MINISTRY OF FINANCE

Evaluation of DSGE model KOOMA with a sign restricted Structural VAR model

Publications of the Ministry of Finance – 2019:62

Discussion papers

Publications of the Ministry of Finance 2019:62

Evaluation of DSGE model KOOMA with a sign restricted Structural VAR model

Oliver Snellman

Ministry of Finance Helsinki 2019

Ministry of Finance

ISBN PDF: 978-952-367-044-0

Helsinki 2019

Kuvailulehti

Julkaisija	Valtiovarainministeriö		11.10.2019
Tekijät	Oliver Snellman		
Julkaisun nimi	DSGE malli KOOMAn arviointi merkkirajoitetulla Rakenteellisella VAR mallilla		
Julkaisusarjan nimi ja numero	Valtiovarainministeriön julkaisuja 2019:62		
Diaari/hankenumero		**Teema**	Keskustelualoitteet
ISBN PDF	978-952-367-044-0	**ISSN PDF**	1797-9714
URN-osoite	http://urn.fi/URN:ISBN:978-952-367-044-0		
Sivumäärä	65	**Kieli**	englanti
Asiasanat	merkkirajoitteet, DSG, SVAR, evaluaatio, Suomi		

Tiivistelmä

Tässä tutkimuksessa arvioidaan Valtiovarainministeriön teoreettisen DSGE malli KOOMAn kalibrointia, käyttäen empiiristä merkkirajoitteilla tunnistettua SVAR mallia. SVAR mallin tilastollisesti merkitseviä impulssivastefunktioita, jotka ovat myös johdonmukaisia riippumatta mallinnusvalinnoista, verrataan vastaaviin KOOMAn impulssivastefunktioihin. Tutkimuksen mukaan molempien mallien impulssivasteilla on yleisesti sama etumerkki, mutta vasteiden suuruuksissa ja kestoissa on eroja.

Kustantaja	Valtiovarainministeriö
Julkaisun myynti/jakaja	Sähköinen versio: julkaisut.valtioneuvosto.fi Julkaisumyynti: vnjulkaisumyynti.fi

Presentationsblad

Utgivare	Finansministeriet		11.10.2019
Författare	Oliver Snellman		
Publikationens titel	Utvärdering av DSGE modell KOOMA med hjälp av den teckenbegränsade strukturella SVAR-modellen		
Publikationsseriens namn och nummer	Finansministeriets publikationer 2019:62		
Diarie- /projektnummer		**Tema**	Diskussionsinitiativ
ISBN PDF	978-952-367-044-0	**ISSN PDF**	1797-9714
URN-adress	http://urn.fi/URN:ISBN:978-952-367-044-0		
Sidantal	65	**Språk**	engelska
Nyckelord	teckenbegränsningar, DSGE, SVAR, evaluering, Finland		

Referat

I denna forskning utvärderas kalibreringen av finansministeriets teoretiska DSGE-modell KOOMA. Utvärderingen utförs med hjälp av den empiriska SVAR-modell som identifierats med teckenbegränsningar. SVAR-modellens statistiskt betydande impulsresponsfunktioner som är konsekventa oberoende av beskrivningsvalen, jämförs med motsvarande impulsresponsfunktioner i KOOMA. Forskningen visar att båda modellers impulsrespons för det mesta har samma förtecken, men att responsen varierar till effekten och varaktigheten.

Förläggare	Finansministeriet
Beställningar/ distribution	Elektronisk version: julkaisut.valtioneuvosto.fi Beställningar: vnjulkaisumyynti.fi

Description sheet

Published by	Ministry of Finance		11 October 2019
Authors	Oliver Snellman		
Title of publication	Evaluation of DSGE model KOOMA with a sign restricted Structural VAR model		
Series and publication number	Publications of the Ministry of Finance 2019:62		
Register number		**Subject**	Discussion papers
ISBN PDF	978-952-367-044-0	**ISSN PDF**	1797-9714
Website address URN	http://urn.fi/URN:ISBN:978-952-367-044-0		
Pages	65	**Language**	English
Keywords	sign restrictions; DSGE; SVAR; evaluation; Finland		

Abstract

The aim of this study is to evaluate the calibration of DSGE model KOOMA of the Ministry of Finance with a SVAR model, which is identified with sign restrictions. I compare impulse response functions from the SVAR model, which are found statistically significant and robust to changes in model specifications, to the equivalent impulse response functions from KOOMA. The findings suggest, that KOOMA generally produce impulse responses with same signs as the SVAR model, but there are some differences in the magnitudes and persistence of the responses.

Publisher	Ministry of Finance
Publication sales/ Distributed by	Online version: julkaisut.valtioneuvosto.fi Publication sales: vnjulkaisumyynti.fi

Content

Abstract 1

1 Literature review 2

2 Theory of DSGE and SVAR models 5

2.1 DSGE modelling 5

2.2 SVAR modelling 7

2.3 Identification of SVAR models 9

3 The empirical SVAR model 16

4 SVAR responses – analysis of significance and robustness 21

5 Comparison of KOOMA and SVAR models 27

6 Conclusions 32

Evaluation of DSGE model KOOMA with a sign restricted Structural VAR model

Oliver Snellman *

Abstract

The aim of this study is to evaluate the calibration of DSGE model KOOMA of the Ministry of Finance with a SVAR model, which is identified with sign restrictions. I compare impulse response functions from the SVAR model, which are found statistically significant and robust to changes in model specifications, to the equivalent impulse response functions from KOOMA. The findings suggest, that KOOMA generally produce impulse responses with same signs as the SVAR model, but there are some differences in the magnitudes and persistence of the responses.

Keywords: sign restrictions; DSGE; SVAR; evaluation; Finland
JEL Classification: C52; E37; O52

*Ministry of Finance, Snellmaninkatu 1 A, Helsinki, P.O. Box 28, FI-00023 GOVERNMENT. Tel. +358 44 231 0677, E-mail: oliver.snellman@gmail.com

The opinions in this paper are those of the author and do not necessarily reflect the opinions of the Ministry of Finance of Finland nor the Government. I am thankful for the assistance of Jukka Railavo, the head of Macro Forecasting, in operating with and understanding the functioning of KOOMA. I also wish to thank my academic adviser Professor Antti Ripatti from the University of Helsinki for useful comments, and Mikko Sariola from the Bank of Finland for providing help. I am grateful for lecturers of courses in macroeconometrics in University of Helsinki for high quality methodological education. Empirical analysis on SVAR and DSGE models was conducted using the IRIS toolbox in Matlab and visualizations produced by R. Possible errors and shortcomings are naturally mine.

Introduction

It is popular in academic and applied research to study macroeconomic phenomena using Structural Vector Autoregressive (SVAR) and Dynamic Stochastic General Equilibrium (DSGE) models. These models can be used to test macroeconomic theories with data, to support in forecasting, to conduct policy and business cycle analysis, and to analyze past events. In Finland, SVAR models have been used for example to decompose the drivers behind two recent depressions (Gulan et al., 2014) and to study the size of the fiscal multiplier effect (Lehmus, 2014). Many Central Banks and other national authorities have developed their own DSGE models, including the Finnish Ministry of Finance and the Bank of Finland, models "KOOMA" and "Aino 2", respectively.

SVAR and DSGE models are comparable to each other under certain conditions, which allows them to be used as substitutes or complements. DSGE models are derived from economic theory and they typically have more a priori structure than SVAR models, and they attempt to paint a broader picture of the economy as a whole. Data driven SVAR models on the other hand are more flexible and suitable for analyzing empirical data. Theory latent assumptions and restrictions can be extracted from a DSGE model to aid identification of structural shocks in a SVAR model, and the parametrization of a DSGE model can be obtained or validated with data by using SVAR models. Both models do, however, require committing to assumptions, which might not all be reasonable.

The aim of this study is to evaluate the calibration of KOOMA on Finnish data. I use time series of six observable variables from KOOMA to identify four of KOOMA's structural shocks from data with a sign restricted SVAR model. Then I compare the impulse responses from KOOMA to the equivalent significant and robust impulse responses from the SVAR model. The statistical significance of the responses is assessed with Bayesian credible intervals and the robustness of the significant responses is analyzed by extensively varying the specifications of the model and variables.

The results indicate, that the signs of the responses are generally the same in both models. The magnitudes of responses to all shocks are generally similar in both models on output, larger in KOOMA with respect to hours, prices and wages, and larger in the SVAR model on responses of exports. The persistence of responses are generally similar on output, hours and exports, and larger in KOOMA on prices and wages. The SVAR results, however, are not fully conclusive due to sampling and model related uncertainty.

This study is structured as follows: Section 1 gives an overview of the literature on the connection between DSGE and SVAR models. Section 2 provides a technical introduction to DSGE and SVAR modelling, focusing on the problem of identification in SVAR models and discussing pros and cons of using sign restrictions. The

SVAR model used in this study is outlined in Section 3. Impulse responses from the SVAR model are presented in Section 4 with robustness analysis. The comparison of KOOMA to the SVAR model is conducted in Section 5, with Section 6 concluding. All of the results from both models and all of the robustness checks can be found in the Appendix.

1 Literature review

DSGE and SVAR models are attempts to answer the famous sceptical epistemic argument presented by Nobel Laureate Robert Lucas. According to the Lucas critique, it is not sensible to estimate the impacts of a change in policy environment on the economy, based on historical and highly aggregated data, as this policy change can also alter the dynamics of the underlying data generating processes (Lucas, 1976). DSGE models attempt to solve the Lucas Critique by rooting into microeconomic theory, i.e. by taking the constrained optimization problems of agents and firms as a starting point, and constructing macroeconomic features on top of this micro-foundation. They provide a theoretically coherent framework with an empirical interface to systematically study economic effects, which cannot be meaningfully isolated (Fernández-Villaverde et al., 2016). SVAR models also try to identify structural features from time series data. More specifically, the parameters in these models are defined as structural, if they are invariant to the class of policy interventions which are being analyzed (Fernández-Villaverde and Rubio-Ramírez, 2008). It is, however, not clear how well these models succeed in answering the Lucas critique. More on SVAR modelling in Subsection 2.3.

DSGE models

The first micro-foundational DSGE model was created by Kydland and Prescott (1982). A new push came from Smets and Wouters (2003), which attempted to find parameter values via Bayesian estimation instead of calibration for a model which includes the whole Euro-area. Two main approaches to DSGE modelling are Real Business Cycle (RBC) models focusing in real variables and technology shocks, extending the tradition of neoclassical growth modelling, and New Keynesian monetary models with real and nominal frictions and a monopolistic supply sector. Most modern large DSGE models are syntheses of features from both approaches.

Critique of DSGE modelling has been issued for example on the use of exogenous shocks (Romer, 2016), the lack of justification for tight priors in estimating DSGE models (Blanchard, 2018), use of representative agents (Solow, 2010), too simplistic micro foundations (Stiglitz, 2018) and the equilibrium framework (Vaughn, 2013;

Hendry and Muellbauer, 2018). Also the validity of conventional distributional assumptions behind many econometrical models has been questioned (Taleb, 2007; Mandelbrot, 2008).

There is a constant effort to make DSGE models more realistic and versatile, for example with heterogeneous agents, non-standard utility functions and time varying volatility (Fernández-Villaverde et al., 2016). Improvement in computing power also permits solving models with more complicated forms and features. Arguments in defence of DSGE modelling can be found for example in Christiano et al. (2018).

A central question related to DSGE models concerns how they should be parameterized and evaluated. Suggested approaches include calibration based on institutional knowledge, estimation based on Bayesian likelihoods or generalized method of moments, and minimizing the distances between DSGE and SVAR impulse responses (An and Schorfheide, 2007). The last method is also motivating the approach in this study.

Connection between DSGE and SVAR models

DSGE and SVAR models are natural econometric counterparts in the sense, that under certain conditions a solved DSGE model, or a subsection of it, can be represented as a (Structural) VAR model of finite order. In that case the structural shocks and their impulse responses from suitably identified SVAR model can be meaningfully compared to those of a DSGE model (Ravenna, 2007).

A solved DSGE model in State-Space form, Equations 1 and 2, can also be written as a VARMA(p,q) model, if there are equal number of state variables, observable variables and shocks. For a VARMA model to have a VAR(∞) representation, invertibility condition[1] must hold (Fernández-Villaverde et al., 2007). Non-invertibility arises fundamentally from the missing information problem, where important variables or factors are not present in the model (Sims, 2012*b*). Finally, a DSGE model can have a VAR(p) representation of finite order[2], in some cases even VAR(1), for observable variables with same structural shocks (Kilian and Lutkepohl, 2017, Chapter 6.2).

Typically only a subset of variables from the DSGE model are included in the corresponding VAR model, namely from the set of observable variables, which have time series data available. Even if the DSGE model would have finite order VAR representation with all of its variables included, this might not be the case for the VAR based on only observable variables. If the DSGE model only admits a VAR(∞) form, any finite order truncation might be inadequate in approximating the real

[1] The eigenvalues of $A - BD^{-1}C$ are strictly less than unity in absolute value, where the matrices A, B, C and D are from the State Space representation of a solved DSGE model, Equations 1 and 2 (Fernández-Villaverde et al., 2007).

[2] DSGE model has a VAR(p) form if all of the eigenvalues of $A - BD^{-1}C$ are zero. For a typical State Space system this entails VAR(1) form.

model, and can also cause identification bias due to structural shocks being based on wrong VAR coefficients (Ravenna, 2007). However, in some cases VAR(∞) can be approximated reasonably well with VAR(p) (Sims, 2012*a*; Pagan and Robinson, 2016).

Evaluation of DSGE models with SVAR models

Many studies have used the SVAR impulse responses to evaluate DSGE models with data. Gali (1999) uses a SVAR model with long run restrictions to compare the RBC and New Keynesian DSGE models, which describe opposite behaviour of hours worked in response to technology shock. The study finds support for the New Keynesian DSGE model, with hours declining in the empirical model after technology shock. The behaviour of hours in response to technology shock is discussed more thoroughly in Section 3. Peersman and Straub (2006) also study the validity of some features of New Keynesian DSGE models. They equip a SVAR model with all applicable sign restrictions suggested by a New Keynesian DSGE model, and compare the resulting impulse responses and forecast error variances to those, calculated from a SVAR model with minimal necessary sign restrictions needed to identify the shocks on interest. They find for example that the model with heavier sign restrictions underestimates the impact of technology shock on output. In similar manner, Canova and Paustian (2011) finds sign restrictions which are robust across a class of DSGE models, identifies a SVAR model with them, and compares the impulse responses which were left unrestricted in SVAR model to the corresponding impulse responses in DSGE models, finding that the same shocks are recovered reasonably well.

Christiano et al. (2007) examine, how well the structural shocks of DSGE models can be recovered by SVAR models in practice. They simulate datasets with different DSGE models, and try to identify the structural shocks from the simulated data with SVAR models using long and short run restrictions. They find that SVAR models with especially short run restrictions accomplished the task well, and also the SVAR confidence intervals reflected the sampling uncertainty accurately. Fry and Pagan (2011), on the other hand, use empirical data to find maximum likelihood estimates for parameters of a small New Keynesian DSGE model, and obtain a sign restricted SVAR model based on a three variable VAR(1) model. They find large differences in the impulse responses.

SVAR impulse responses have also been used to estimate DSGE models. The approach in Christiano et al. (2005) to fully parametrize their partially calibrated New Keynesian model is to first use a SVAR model to obtain impulse responses to monetary policy shock, and then find the values for the rest of the parameters by minimizing the distance between the IRFs from both models. Similarly, Fève et al. (2010) use long run restrictions to identify impulse responses to disinflation shocks in a SVAR model. They then use the IRFs to estimate a partially calibrated medium-sized DSGE model by minimizing the distances between the responses, and use the DSGE model to study the disinflation shocks with counterfactual scenarios. Fève et al. argue, that SVAR models are better suited to study transitory changes in monetary policy, and permanent changes for example in inflation can be approached better using a DSGE model.

Del Negro and Schorfheide (2004) suggests estimating a BVAR model, coined as DSGE-VAR(λ), with priors derived from a DSGE model. The hyperparameter λ controls the weight given to DSGE model relative to data in estimation. They continue developing their approach in (Del Negro and Schorfheide, 2006) and (Del Negro et al., 2007), which can be used either to analyze the DSGE model, or as an interface to forecast on the basis of the DSGE model. Consolo et al. (2009) suggest using principal components to create a factor augmented DSGE-FAVAR instead, which they found to perform better.

Brüggemann and Kascha (2017) study whether a VARMA model can discriminate between DSGE models better than a SVAR model, but find neither performing well. Also Leeper et al. (2013) use a VARMA model to incorporate informational variables to the SVAR model when replicating a well known study by Blanchard and Perotti (2002). They argue that the correct shocks were only obtainable with the added information.

2 Theory of DSGE and SVAR models

Subsection 2.1 provides a non-technical intro to the idea of DSGE modelling. Subsection 2.2 presents the basic technical functioning of VAR and SVAR models, while Subsection 2.3 discusses different ways to identify SVAR models on the basis of VAR models, with an emphasis on sign restrictions.

2.1 DSGE modelling

DSGE models are dynamic in the sense that they evolve in time, stochastic as they have inbuilt randomness, general as they try to model all markets of the economy, and have equilibrium behaviour, which is based on the assumption that markets are driven towards balance by the forces of supply and demand. DSGE models can be used to study, how a system with certain characteristics converges back to its steady state after an exogenous shock, policy change or from cyclical position of boom or bust. This Section is largely based on (Fernández-Villaverde et al., 2016).

DSGE models are derived from microeconomic intertemporal constrained optimization problems for consumers, firms and other entities. To find the solution of a

DSGE model, these optimality conditions must first be transformed into a functional equation problem for example by using Euler equation or conditional expectations. The result is a set of non-linear stochastic difference equations, which are meant to characterize the functioning of the economy of interest, Finland in the case of KOOMA. The model variables are categorized as control and state variables, where the control variables are observable and forward looking, like the Gross Domestic Product, and the state variables are typically not observable to econometrician and are considered to be predetermined or backwards looking, like the technological development or the capital stock. It is, however, not always obvious which variables should be treated as observables.

The solution of the model refers to a set of decision functions for the control and state variables, which describe the behaviour of the system. These decision functions consist of state variables, stochastic shocks and unknown parameters. As DSGE models don't typically have a closed form solution for their equilibrium dynamics, the solution is approximated numerically using perturbation theory. A deterministic steady state of the system, describing the equilibrium conditions, can typically be found analytically by suppressing the stochastic components to zero. Taylor or log-linear approximations are then built around the steady state to approximate how the system behaves around the equilibrium conditions. Typically first and second order perturbations are used, but higher order perturbations are needed when approximating non-linear behaviour and stochastic volatility (Levintal, 2017). A stable solution can be found by transforming this approximated system into a State-Space representation using for example generalized Schur decomposition.

Solved DSGE model in (simplified) State Space form

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{v}_t \tag{1}$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{w}_t \tag{2}$$

consists of measurement Equation 1 for the observable variables in vector $\mathbf{y}_t$, and state transition Equation 2 for the control variables in vector $\mathbf{x}_{t+1}$, with corresponding stochastic shocks $\mathbf{w}_t$ and $\mathbf{v}_t$. Matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are functions of the parameters of the DSGE model. (Fernández-Villaverde et al., 2007)

The model parameters in matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ can be assigned values based on theory and institutional knowledge by calibration, or by data via numerical Bayesian estimation using Markov Chain Monte Carlo (MCMC) algorithms, called estimation. The parameters governing technological development and preferences are called deep parameters of the model and are assumed to be invariant to policy changes. Values for the unobserved state variables can be searched numerically using Kalman filter. (An and Schorfheide, 2007)

Linear approximation is assumed to be adequate for studying the behaviour of the system around the steady state. The upside of linear approximation is that it provides intuitive results, for example log-linearization gives the approximate percentage deviation from the steady state. It is also possible to use non-linear approximations, but this makes the analysis complex without necessarily increasing usefulness (Levintal, 2017).

KOOMA

The KOOMA model is a calibrated New Keynesian DSGE model, describing a small open economy. This means that the model economy is a price taker in foreign trade. The model includes two types of households, public sector, foreign trade, and sectors for intermediate and final goods. The model is log-linearized with first order Taylor series. More information can be found in the forthcoming manuscript of the KOOMA model, which is not yet made public by the time this study was published.

Shocks of interest

The shock processes represent exogenous changes in the DSGE model's environment. The four shocks in KOOMA, which are relevant for this study, are AR(1) processes with i.i.d. innovations. **Technology shock**, $\varepsilon_t^a = \rho^a \varepsilon_{t-1}^a + \zeta_t^a$, refers to changes in the productivity, with the autoregressive parameter calibrated as $\rho^a = 0.9$. **Labour supply shock**, $\varepsilon_t^L = \rho^L \varepsilon_{t-1}^L + \zeta_t^L$, alters the supply of labour at given wage level, with calibration of $\rho^L = 0.37$. **Domestic demand shock**, $\varepsilon_t^C = \rho^C \varepsilon_{t-1}^C + \zeta_t^C$, changes the households' willingness to consume at given price level, with $\rho^C = 0.8$. **The external demand shock**, $\chi_t^* = \rho^{\chi^*} \chi_{t-1}^* + \zeta_t^{\chi^*}$ with $\rho^{\chi^*} = 0.73$, represents changes in the global aggregate demand for imports, which is closely related to the country specific export demand shock identified in the SVAR model. The first two shocks are related to the supply side and latter two to the demand side of the economy.

2.2 SVAR modelling

This Subsection is based on Kilian and Lutkepohl (2017).

VAR model

Vector Autoregression model VAR(p) displays the linear multivariate autoregressive behaviour of a system of K variables in matrix notation, using lagged values of these variables from past p time periods as explanatory variables (omitting a constant)

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t \tag{3}$$

where y_t and y_{t-i} with $i = 1, \ldots, p$ are Kx1 vectors holding values of the K variables for contemporaneous and lagged time periods, respectively, A_i are KxK matrices of coefficients describing the influence of lagged variables on their contemporary values, and u_t is a Kx1 residual vector of i.i.d. disturbances, with a covariance matrix (apostrophe denoting for transpose)

$$\mathbb{E}(u_t u_t') = \Sigma_u \tag{4}$$

The parameters in the coefficient matrices A_i can be consistently estimated with equation-by-equation Ordinary Least Squares (OLS).

If the process is weakly stationary, i.e. roots of the characteristic polynomial lie within the unit circle,

$$det(I_k - A_1 z - \cdots - A_p z^p) \neq 0, \; |z| \leq \; 1 \; (z \in \mathbb{C}) \tag{5}$$

the VAR(p) model has a vector moving average VMA(∞) representation

$$y_t = \sum_{j=0}^{\infty} \phi_j u_{t-j} \tag{6}$$

where ϕ_j matrices are obtained recursively from $\phi_j = \Sigma_{i=1}^{x} \phi_{j-i} A_i$ with $\phi_o = I_k$, and $x = j$ when $j < p$ and $x = p$ when $j \geq p$. The weakly stable process can be thought of as being driven by the error terms and we can study the propagation of the j^{th} error with Impulse Response Functions (IRF)

$$IRF_{.j}(h) = \frac{\partial y_{t+h}}{\partial u_{jt}}, h = 0, 1, 2, \ldots \tag{7}$$

which are expressed in the j^{th} column of ϕ_h. In macroeconomics the standard VAR model presented above is referred to as having a reduced form, as the model more likely summarizes data instead of providing an interpretation of its underlying dynamic process. This is because the elements of the residual vector are typically cross-correlated. Hence, it is not sensible to analyze the impulse responses of individual reduced form errors in isolation, as their movements tend to be synchronized.

Structural VAR model

We are interested in finding a VAR model, which displays the behaviour of chosen variables as driven by uncorrelated structural errors, which can be interpreted as economic shocks. To find this structural form, we write the same reduced form VAR(p) model in another representation

$$y_t = B_0^{-1} B_1 y_{t-1} + \cdots + B_0^{-1} B_p y_{t-p} + B_0^{-1} \varepsilon_t \tag{8}$$

with decompositions $B_0^{-1} B_i = A_i$ and $B_0^{-1} \varepsilon_t = u_t$. Identifying B_0^{-1} matrix and multiplying both sides by its inverse, B_0, results in Structural VAR (SVAR) model

$$B_0 y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t \tag{9}$$

where the structural shocks, ε_t, are linear combinations of reduced form residuals

$$\varepsilon_t = B_0 u_t \tag{10}$$

and are orthogonal (uncorrelated) by construction

$$\mathbb{E}(\varepsilon_t \varepsilon_t') \equiv diag(\Sigma_\varepsilon) \tag{11}$$

The key component in finding the structural model is to identify the B_0^{-1} matrix describing the transmission of structural shocks on model's variables. The structural shocks are typically assumed to have a unit variance, $diag(\Sigma_\varepsilon) = I_K$[3] and the B_0^{-1} matrix can be obtained from the reduced form VAR residuals

$$\mathbb{E}(u_t u_t') = \mathbb{E}(B_0^{-1}\varepsilon_t \varepsilon_t' B_0^{-1\prime}) = B_0^{-1}\mathbb{E}(\varepsilon_t \varepsilon_t') B_0^{-1\prime} = B_0^{-1} I_K B_0^{-1\prime} = B_0^{-1} B_0^{-1\prime} \tag{12}$$

Different techniques to recover the B_0^{-1} matrix from the VAR residuals will be discussed in the next Section.

When B_0^{-1} is identified, the VMA representation of weakly stable SVAR model is obtained by substituting $u_t = B_0^{-1}\varepsilon_t$ in the reduced form VMA representation defined in Equation 6

$$y_t = \sum_{j=0}^{\infty} \phi_j u_{t-j} = \sum_{j=0}^{\infty} \phi_j B_0^{-1} \varepsilon_{t-j} = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} \tag{13}$$

where $\theta_t = \phi_j B_0^{-1}$, which the structural IRF is based on.

2.3 Identification of SVAR models

The challenge of identification is, that there are K^2 parameters in B_0^{-1}, more than there are unique equations, hence the structural model cannot be identified solely from the data. There are different methods that can be used to find orthogonal errors from the residuals of reduced form VAR model. These orthogonal errors can then be addressed with an economic interpretation, resulting in locally and uniquely

[3]This assumption is called B–normalization, and it gives the impulse responses to unit size shocks. It is also possible to use A–normalization, where the diagonal elements of structural error covariance matrix is left unrestricted and the diagonal elements of are restricted to unity instead. A–normalization expresses the sizes or relevance of structural shocks relative to each other in the identified model. AB–normalization is a general version, where both relative sizes and responses to unit shocks can be obtained, but with the cost of requiring more assumptions to enable identification.

identified structural VAR model. Table 1 depicts the most common identification techniques, which include posing restrictions on the model parameters, taking advantage of statistical properties of the data and using a Bayesian approach. It is noteworthy, that different techniques usually lead to identification of different orthogonal errors as structural shocks, and hence result in different structural models. It is not self-evident which one of these (infinite) models, if any, is the correct one. The question of how to identify structural models was deemed the most important problem in empirical macroeconomics by Lucas Jr. and Sargent (1979).

Parametric	Recursive	Non-Recursive
Contemporaneous	Cholesky decomposition with A or B normalization	With A, B or AB normalization
Long (and short) run	BQ (Cholesky) decomposition on long run impact matrix	Long and/or short run restrictions on SVEC model
Statistical properties	Paradigm shift in volatility, heteroscedasticity, non-Gaussian or Markov switching reduced form residual distribution.	
Bayesian approach	Bayesian VAR (typically with Minnesota prior), **Sign Restrictions**	

Table 1: Taxonomy of SVAR identification methods. Author's sketch.

Parametric restrictions

Parametric restrictions can be imposed recursively or non-recursively on the structural model parameters, which define the contemporaneous and/or long run impacts of the structural shocks. In order to exactly identify the contemporaneous impact matrix B_0^{-1}, the necessary order condition must hold; the number of restrictions issued on parameters of B_0^{-1} must be equal to its unique equations, which is $K^2 - K(K+1)/2 = K(K-1)/2$ by the symmetry of covariance matrices. Additionally, necessary and sufficient rank condition must hold for identification, requiring B_0^{-1} to have full rank. Applying Cholesky decomposition on the residual covariance matrix Σ_u (Sims, 1980), or on the long run impacts of the shocks (Blanchard and Quah, 1989), results in a recursive lower triangular matrix satisfying the number of restrictions needed. The free parameters can be estimated by Maximum Likelihood method with additional distributional assumption for the error term. It is also possible to non-recursively restrict B_0^{-1}, or the short and long run impact matrices in Structural Vector Error Correction (SVEC) representation. The use of long run restrictions require the presence of unit root processes, which shocks could have permanent impact on.

The challenge with parametric restrictions is, that they require committing to strong a priori assumptions about the relations between the structural shocks and the variables of the model, which might be hard to justify and don't seem compatible with results from some theoretical models. Also the number of parameters, and thereby the number of necessary restrictions needed, grows exponentially in relation to the number of variables added to the model. Over- and under identifying restrictions can be tested with Likelihood ratio test, but it is not possible to pre-emptively select the correct parametrically restricted structural model with statistical tests.

Statistical properties

Data driven approach avoids the troubles of cherry picking parametric restrictions, by using statistical properties of the data for identification. Structural shocks can be identified, if there is a suitable structural break in their volatility in data, or if they exert heteroskedastic behaviour (Rigobon, 2003; Lanne and Lütkepohl, 2008). It is also possible in some cases to statistically identify the structural shocks if the reduced form residual distribution is assumed to be non-Gaussian, or the error variance is modelled as a Markov switching process, where the structural shocks remain orthogonal in different regimes (Lanne et al., 2010; Sims and Zha, 2006).

Missing information problem

One of the central challenges with SVAR modelling is the missing information problem, where the starting point VAR model lacks some relevant variables or information, which are contributing to the functioning of the phenomenon under consideration. Sometimes the missing information is hard to quantify to begin with, such as expectations about future events. The structural shocks cannot be correctly identified under missing information, as all of the components comprising the shocks are not present in the data. Adding more explanatory variables to the model brings forth other challenges especially with parametric identification strategies, as mentioned above, as the number of restrictions required grows exponentially in relation to model variables.

The Bayesian VAR (BVAR) model allows increasing the number of variables in the model without enormously increasing the number of restrictions needed for identification (Doan et al., 1984). BVAR framework also avoids the need to commit to binding parametric restrictions, by instead assigning informative prior distributions on the parameters, typically the Minnesota prior scheme. Factor Augmented VAR (FAVAR) attempts to tackle this type of missing information problems by including principal components from large pools of potentially relevant variables as explanatory variables in the VAR model (Bernanke et al., 2005).

Sign Restrictions

Lately it has become popular to use sign restrictions, a Bayesian approach in spirit. Instead of concentrating on the parameters of the model, the signs of the impulse responses of the structural model are restricted. Sign restrictions have been used as means of identifying SVAR models since Faust (1998), and the methods have been developed further by Canova and Nicolo (2002), Uhlig (2005) and Rubio-Ramírez et al. (2010), among others. Sign restrictions offer a softer approach to identification than strict parametric restrictions. In the beginning, sign restrictions were mainly used to identify single shocks, particularly the monetary policy shock. Since then they have been applied to identify other shocks as well, and to identify multiple shocks at a time.

Identification via sign restrictions requires conditional simulation of potential models, of which are accepted for further considerations based on fulfilling the prescribed sign restrictions. The rationale behind this method is, that the possible structural features generating the data should also translate to the conditionally simulated models. By concentrating on a chosen subset of simulated models based on the sign restrictions, some interesting features can be isolated for further examination. The need and justification for using simulations also partly derives from the scarcity of macroeconomic time series data.

As we have narrowed our search down to those simulated models that fit our a priori description, the signs of the shocks' impulse responses are not anymore informative about the data. However, the persistence and magnitude of these impulse responses can be revealing, as well as the other variables which didn't have any restrictions posed on their responses to the shocks of interest. By simulating large number of models, we can gather a sample to approximate the posterior distributions of IRFs from the set of models which satisfy the chosen sign restrictions.

The procedure begins with estimation of reduced form VAR(p) model to obtain coefficient matrices A_i and the residual covariance matrix Σ_u from data. New model is then created by drawing new coefficient matrices from Normal distribution and a residual covariance matrix from inverse-Wishart distribution, which are both conditional on the estimated VAR coefficient and covariance matrices. Orthogonal errors are identified from the conditionally simulated model, typically by conducting a Cholesky decomposition on residual covariance matrix $u_t u_t' = \hat{\Omega} = PP'$ with lower diagonal P to get $\varepsilon_t = u_t P^{-1}$, and $\mathbb{E}(\varepsilon_t \varepsilon_t') = I_K$. Recursive identification is used here just to get starting point orthogonal shocks easily, but any scheme would suffice. A linear transformation is then applied on these errors with a weighting matrix Q to produce a candidate draw from the set of all structural models. The candidate model has a contemporaneous impact matrix $H = QP^{-1}$ and shocks $\hat{\varepsilon_t} = Q\varepsilon_t$. The impulse response functions of the candidate model are evaluated, and if they fulfil the sign restrictions, they are considered to be draws from the desired posterior distributions.

In order to preserve orthogonality of the candidate shocks $\hat{\varepsilon}_t$, Q is required to be a square matrix with orthogonal columns of unit vectors, so that $Q'Q = QQ' = I_K$. Then $\mathbb{E}(\hat{\varepsilon}_t \hat{\varepsilon}_t{}') = \mathbb{E}(Q\varepsilon_t\varepsilon_t' Q') = Q\mathbb{E}(\varepsilon_t\varepsilon_t')Q' = I_K$, fulfilling the orthogonality requirement. Suitable unique Q matrices can be found with QR decomposition of a random invertible matrix X, which have elements independently drawn from the standard normal distribution, and where the R matrix is upper triangular with positive diagonal elements. The QR decompositions are conducted using Householder's transformations in IRIS.

It is possible to issue the sign restrictions on impulse responses for one or more periods, or to issue the restrictions on the sums of responses from some interval, as in (Sariola, 2015). The algorithm can be made more efficient by checking whether multiplication of columns of Q by -1 produces an accepted model. Likewise, columns of impact matrix H can be switched to match the sign restrictions, making the ordering of variables also irrelevant with respect to the end result. Both procedures produce new structural matrices, which are equally applicable.

This procedure is repeated until the required number of accepted draws have been obtained.

The next step is then to choose a representative model from the set of accepted models. One way to do that would be to take the pointwise medians of the impulse responses across all models. I utilize a different method in this study to guarantee that all of the impulse responses come from the same model. The choosing criterion is, that the representative model must be the closest model to median responses across all IRFs. To find this model we compare the values of impulse response functions of identified shocks on all variables, from all accepted simulated models, at each period on chosen time interval, to the respective median responses. The distances are then squared and summed up.

$$\sum_{i \in I, j \in J, k \in K} [S_{i,j}(k) - M_{i,j}(k)]^2 \tag{14}$$

Here $S_{i,j}(k)$ is the value of the impulse response function of the j^{th} shock on i^{th} variable in k^{th} period, and $M_{i,j}(k)$ stands for the corresponding median response value among all simulated models. In this study, the set of variables, I, includes all model variables, set of shocks, J, includes four identified shocks, and set K includes four first periods after the innovations.

All simulated models are ranked according to the sums of squared distances from their chosen impulse responses to the corresponding means. The model with the smallest sum is chosen as the representative model of the set of all models that fulfil the sign restrictions.

The uncertainty of the impulse responses from the representative model can be expressed by constructing credible intervals of $(1-\alpha)$ %, by depicting the point-wise $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ percentile responses at each period from posterior distributions of each IRF, and by regarding them as the lower and upper bounds, respectively, of the credible intervals. These intervals have a Bayesian interpretation, as they describe the posterior distributions of IRFs, and the values of these bounds can come from different simulated models. Granziera et al. (2018) note, that the large sample numerical equivalence between frequentist confidence sets and Bayesian credible sets doesn't apply on set identified models, and therefore these credible intervals cannot be addresses with an interpretation as approximate confidence intervals.

Challenges with sign restrictions

It is not clear to what extent these simulated models describe the actual data, because simulating large number of models allows for many different kind of shocks to be found. However, the validity of chosen identifying restrictions can't be meaningfully measured alone by the ratio of accepted models to all simulated models either, as the ratio might be impacted by a factor of 10 depending on the efficiency of the algorithm, among other factors. Conversely, too high ratio of accepted models to all simulated models might indicate insufficient restrictions. (Kilian and Lutkepohl, 2017)

Each individually accepted structural shock comes from a different model with another $K-1$ shocks, K denoting the number of variables in the model, and these additional shocks are typically not paid attention to. Lucas (1976) reminds, that altering these other shocks can have an influence on the behaviour of the shock under consideration, so it might not be enough to just concentrate on identified shocks. The identified shock is also only guaranteed to be internally orthogonal to the other shocks in the same model, not to other shocks separately identified by sign restrictions on the basis of the same original VAR model.

If multiple shocks are to be identified using sign restrictions, they must be identified from the same simulated model and therefore be searched simultaneously by inspecting all restrictions on IRFs at once (Uhlig, 2005). This can lead to difficulties as the particular matrix we wish to find might be rare, requiring exhaustive simulations and resulting in extremely low acceptance rate, in some cases in the magnitude of one in millions.

Sign restriction only allows for set identification, deriving from the use of inequality restrictions (Kilian and Lutkepohl, 2017, page 414). There are infinitely many parametrizations for simulated structural models, which can fulfil the same sign restrictions. This derives from the fact that for each identified model, there exists an orthogonal matrix arbitrarily close to an identity matrix, which can be used to multiply the model, with the resulting new model still satisfying the same sign restrictions (Rubio-Ramírez et al., 2010).

It is not clear to what extent the variation in shocks' impulse responses comes from data and model related uncertainty, that is, from uncertainty about the true reduced form VAR covariance matrix and coefficients (or the correct functional form of the model), and from the procedure by which the structural model was identified from them. The credible intervals describe the distribution of simulated models and is influenced by both factors of uncertainty.

In some cases an identified shock might actually be a linear combination of multiple structural shocks, which all fulfil the same restrictions. For example, Uhlig (2005) was worried that money demand shock and monetary policy shock might not be separable by sign restrictions alone. Fry and Pagan (2011) suggest, that sign restrictions should be used together with parametric restrictions. Parametric restrictions can be combined with sign restrictions by applying Givens rotations on the identified H matrix and solving for the angles of rotation resulting in desired restrictions. Any parametric restriction on structural model identified with sign restrictions will lead to over-identification, and hence allow for testing for the validity of the parametric restrictions via likelihood ratio.

Baumeister and Hamilton (2015) stresses, that the Bayesian element is commonly not acknowledged properly in the use of sign restrictions. There typically is an implicit and informative prior distribution on the background, which in some cases influences the analysis even when the number of conditional simulations approaches infinity. In that situation, the researcher is merely studying the prior distribution when conducting identification by sign restrictions. They for example notice, that common algorithms like the popular "RWZ rejection algorithm" proposed in Rubio-Ramírez et al. (2010), only work on specific uniform Haar prior for the Q matrix, but this prior is not uniform in all respects. The procedure used in X=QR decomposition, due to orthonormality of the columns of Q, results in smaller values for the elements of Q when the number of variables in the VAR model, and hence the dimensions in X increase. This results in the prior distribution for the elements of $H = PQ$ to be flat only when $K = 3$, to favour values closer to zero when $K > 3$, and to place more weight for larger deviations of parameter values from zero when $K = 2$.

3 The empirical SVAR model

In this Section I detail out the variables, VAR model, sign restrictions and diagnostics behind the SVAR model used in this study. First I estimate a VAR(2) model with six variables and an intercept by OLS, and then apply the sign restriction based simulation and sorting algorithms detailed in Subsection 2.3.

Reference points for the model used in this study include Peersman and Straub (2009), who study the impact of technology shock on hours worked in the Euro area with a sign restricted SVAR model, and also use the results to analyze the different behaviour of RBC and New Keynesian DSGE models on that regard. Based on their model, Sariola (2015) studies the Swedish business cycle, in companion with and comparison to DSGE model "Ramses 2" of the Bank of Sweden, providing a good starting point for this study.

Variables

The model has six variables: output, total hours worked, inflation (CPI), real hourly wages, real interest rate and exports. The data is obtained from Statistics Finland, except for the interest rate, proxied by the value of 10 year government bond, which was provided by the Treasure of the State. All variables are seasonally adjusted logarithmic Year-on-Year differences, except the interest rate, which is otherwise similar but in absolute values. I specified it in this way to avoid extreme values, caused by the time series approaching zero in recent years. For this reason the impulse responses of interest rate are not interpreted as approximate percentage changes, as is the case with other responses. In order to use exactly the same variables as in KOOMA, I used private consumption prices as a measure of inflation instead of overall consumption prices.

The time series for total hours worked still exerted seasonal variation, although it had already been seasonally adjusted by Statistics Finland. Ahola (2012) provides plausible explanations for this, for example, the method and frequency of inquiring the information about hours worked has changed three times in the past, possibly influencing the statistical properties of each segment. Following Ahola, I removed the seasonal component from hours worked by adjusting the original monthly time series with X-13-ARIMA-SEATS seasonal adjustment software by the US census Bureau (see Sax and Eddelbuettel, 2018) in two parts, 1990 Q1 – 1999 Q4 and 2000 Q1 – 2018 Q4, the cut-off date being the latest methodological change in the inquiry process. I then aggregated the data to quarterly frequency and trimmed it to the same length as the other time series.

Hourly nominal wages are obtained by dividing the aggregate annual wages by the new total hours worked -series. Nominal wages and nominal interest rate are divided by prices to obtain their real counterparts, and real wages -series is then multiplied by 100 to scale it back. Real interest rate, however, is not scaled accordingly, as it is not specified in logarithmic differences like the other variables, to diminish its impact on the sorting algorithm presented in Equation 14 selecting the representative model by their summed squared distances to means. I test the impact of multiplying the interest rate by 100 in robustness check 15 in Section 4, with main difference being in the magnitudes of the responses of interest rate, which was to be expected.

The time series for all variables run from 1999 Q1 to 2017 Q4, although data was available from 1990 Q1. This choice was partly motivated by the need to standardize the data with KOOMA, which only has data for all of its variables available from 1999 Q1 onwards. The data has to be the same in both models in order to obtain comparable sizes for unit innovations and behaviour for the shock processes. Also, Finland was under a different monetary policy regime pre 1999, causing potential qualitative changes in the behaviour of the studied phenomena. Full data is used in robustness check 8, producing some differences in individual IRFs.

Structural shocks

The structural shocks to be identified with these variables are technology shock, labour supply shock, and domestic and export demand shocks. Export demand is the Finnish share of external demand. Even though export demand can also be influenced by changes in exchange rates and national inflation, it is strongly correlated with external demand, the comparable shock in KOOMA. As there are fewer shocks than variables, the model is only partially identified. The two unidentified shocks do fulfil the orthogonality requirement, but do not produce significant responses, plotted in the Appendix. These shocks are not addressed with economic interpretation, as they are linear combinations of the rest of the potential structural shocks influencing the variables.

Sign restrictions

The restricted impulse response functions are required to have the chosen sign for the sum of their values calculated from four period interval after the innovation. This is a more flexible restriction than requiring the IRFs to have the chosen sign on every four period, as only the net impact is restricted. This allows the data to indicate, if it doesn't naturally take the form of strictly positive or negative responses. I restrict only the initial responses in robustness check 2, with mostly similar results.

In differentiating shocks with sign restrictions, opposite restrictions must be issued on the responses of two shocks on the same variables. Specific restrictions used in this study can be found in Table 2. Supply shocks are separated from demand shocks with an assumption of their opposite impact on inflation. Demand shocks are separated by issuing opposite impacts on exports. Technology shock is distinguished from labour supply shock by their differing impact on total hours worked, as suggested by the New Keynesian DSGE literature.

There is an ongoing debate on the impact of technology shock on hours worked. RBC-models in line with Kydland and Prescott (1982) tend to advocate for positive impact, whereas New Keynesian models in line with Gali (1999) suggest that aggregate demand doesn't change immediately due to monetary policy and price rigidities, causing hours to decrease instead. Avoiding taking stances, Peersman and Straub (2009) differentiate the two supply shocks by their opposite impacts on real wages. Although the negative impact of labour supply shock on wages was found robust, the authors note that the impact of technology shock on wages is influenced by the values of model parameters. For example, under high price rigidity the response of real wages can become negative with respect to technology shock, in which case the supply shocks couldn't be adequately separated by restricting their impact on wages. Even though Peersman and Straub found that hours worked increases after technology shock, an opposite impact occurs in one of their robustness checks, when technology shock is identified with long run restrictions. Many other studies have found support for the hypothesis, that technology shock decreases hours worked, for example Kimball et al. (2006) by using growth accounting approach, and Canova et al. (2013) when differentiating technology shocks to neutral and investment specific components. Also KOOMA exerts separable behaviour of supply shocks on hours worked, but not on real wages. Issuing the restrictive choice on hours worked instead of wages is partially motivated by the aim of this study, which is to evaluate the calibration of KOOMA, not the validity of its functional form. Also, I separate the supply shocks by issuing sign restrictions on wages instead of hours in a robustness check 3, which did not affect the responses of hours, but resulted in lower initial response of wages to labour supply shock.

	Output	Hours	Prices	Wages	Interest rate	Exports
Technology shock	+	−	−			
Labour supply shock	+	+	−			
Domestic demand shock	+		+			−
Export demand shock	+		+			+

Table 2: Sign Restrictions used to identify the SVAR model.

In order to compare the size and persistence of impulse responses from SVAR and KOOMA models, the unit size of the shocks must first be standardized similarly. Pagan and Robinson (2016) note, that different initial impulses are a typical cause behind differing responses. The impulse response functions in SVAR model describe typical responses to one standard deviation size unit shocks. To normalize KOOMA's shocks similarly, the data for observable variables is driven through calibrated KOOMA using Kalman Filter to first find values for unobserved state variables, and then to identify a combination of innovations and their sizes, which would have produced the values for the state and measurement variables as a result of the shock processes. The standard deviations of these innovations are then calculated and regarded as the sizes for comparable unit shocks.

Diagnostics

I chose 2 lags for the VAR model as a middle way between suggestions by Akaike and Bayesian information criteria, with more weight given to the latter one, alongside with a constant, see Table 5 in the Appendix. Breuch-Godfrey and Portmanteau tests suggest that there might be some autocorrelation left, but the model does not seem to exert heteroskedastic behaviour. Multivariate Jarque-Bera test doesn't indicate non-normality, nor excess kurtosis or skewness on VAR residuals. The model is stable, as all of the eigenvalues of the companion matrix of the VAR model have modulus less than one, which is equivalent to the stability condition in Equation 5. Hours, prices and wages seem to exert Granger causality on other variables both individually and combined. Johansen's trace test suggests, that there can be up to three cointegration relations among variables. The representative model was chosen among 5000 accepted models fulfilling the sign restrictions, which required the simulation of 13 million structural models. Figure 1 visualizes the ranking of these 5000 models by their score on the loss function of Equation 14. The non-linearity of the curve means, that there are few models really close to or far from the medians, with majority of the models in between. This non-linearity does not seem to be sensitive to the number of simulations.

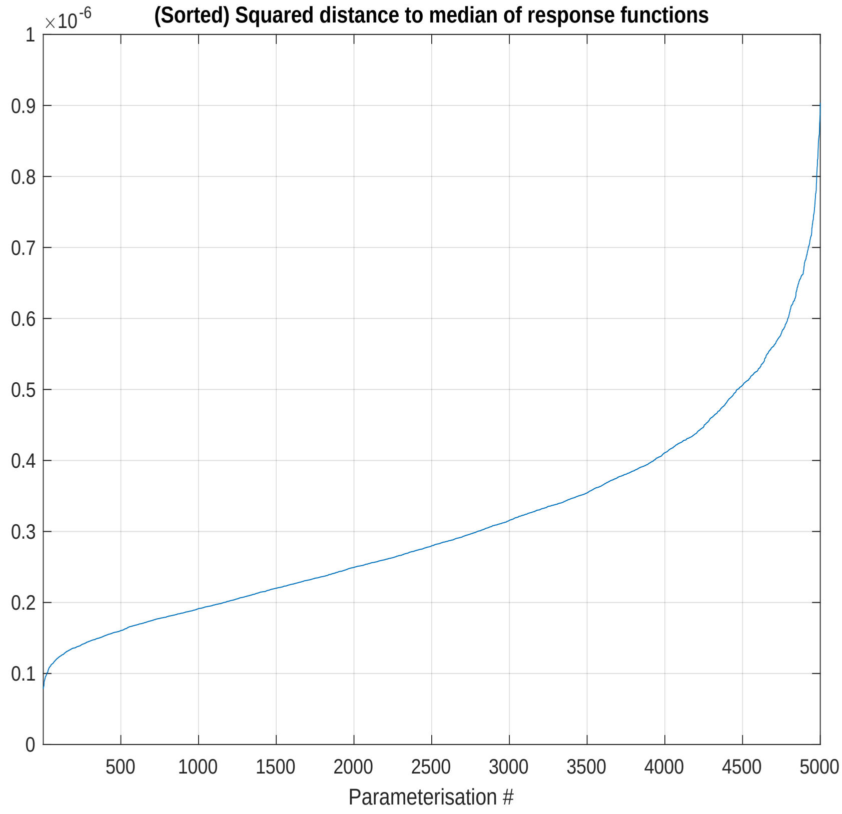

Figure 1: Simulated accepted models sorted by the loss function.

Credible intervals

I use one standard deviation Bayesian credible intervals to asses which responses are significant within the model, and to determine whether the responses from KOOMA and SVAR models differ significantly. Response of a variable is deemed statistically significant in the SVAR model, if zero is not included in the credible interval. The credible intervals of 68 % were constructed by choosing the pointwise 16^{th} and 84^{th} percentile responses at each period from posterior distributions of each IRF. For the sake of practicality, I will refer to these credible intervals as 68 % bands. I chose one standard deviation as the width of the interval instead of larger intervals to leave out the tail events from the analysis, as the tails are likely to be too heavy to be reliably approximated in this manner. Also 90 % bands are plotted in Figure 22 in the Appendix.

4 SVAR responses – analysis of significance and robustness

In this Section I present the IRFs from the SVAR model and analyze their statistical significance and robustness. First I explain how the robustness analysis was conducted and then go through the responses by variable and by shock.

Robustness analysis

To test the robustness of the responses, I vary three parts of the process by which they are acquired: 1) the procedure by which the final SVAR model is obtained on the basis of a VAR model, in robustness checks 1–6, 2) the specifications in estimating the initial VAR model, in robustness checks 7–9, and 3) the variable selection and transformations in robustness checks 10–15. In all of the robustness checks, except in 1^{st} and 4^{th}, I draw new 2500 accepted structural models and choose a representative model among them for the comparison. In robustness checks 7–15, I also estimate the VAR model again with the altered specifications before utilizing the simulation algorithm. Robustness check 5 indicates that the posterior distributions converge fast even with low number of draws, and therefore I lowered the required number of accepted draws from 5000 to 2500 for robustness check models. Summary statistics of the 15 robustness check models is gathered in the Table 3. Description of each robustness check models and all of the individual impulse response comparisons can be found in the Appendix.

An IRF is regarded robust, if most of the responses from different model specifications are contained within the credible intervals of the SVAR model during the periods when the IRF is significant, and if the deviant responses mostly have the same sign. It is not straight forward, how the robustness of a response or lack thereof should be defined. Some of the responses can be expected to lay outside of the bands even if they came from the same model, as the credible interval only covers about two thirds of the expected possible values. I will refer to responses violating the above-mentioned loose condition of robustness as somewhat robust or not robust, based on the specific situation. It is noteworthy, that there is an asymmetry related to the analysis of robustness. If the responses from robustness check models on a significant IRF exceed the credible band on the opposite side from zero, the sign of the response can still significant and robust whereas the magnitude is not.

Model specification	**Akaike**	**Bayesian**	**Simulations (millions)**
The representative model	−54.30	−51.80	6.5**
1. Three best models	∗	∗	∗
2: Sign restrictions only on the first period	∗	∗	4.4
3: Sign restrictions on wages instead of hours	∗	∗	14.1
4: Posterior medians	∗	∗	∗
5: Number of accepted simulations. 2500	∗	∗	6.5
500	∗	∗	1.3
6: Different seed	∗	∗	13.1
7: Different number of lags. 4 lags	−55.12	−50.23	6.0
3 lags	−54.35	−50.66	5.3
8: Data used from 1990 instead of 1999	−51.80	−49.84	12.4
9: Only pre crisis (2008) data	−56.81	−53.24	6.2
Only post crisis	−55.07	−51.64	30.3
10: Trend gap instead of Year on Year	−59.67	−57.17	8.6
11: Different measure for inflation	−55.86	−53.35	6.7
12: Employment instead of hours	−55.19	−52.69	13.4
13: Hours as smoothed by StatFin	−53.71	−51.20	7.8
14: External demand instead of exports	−56.34	−53.70	3.1
15: Real interest rate *100	−45.09	−42.59	31.6

Table 3: Summary of Akaike and Bayesian (Schwarz) information criteria concerning the different VAR specifications and number of simulations needed to produce 2500 accepted structural models.
∗ Same as with the representative model.
** This number is comparable to others, as the representative model required 13 million simulations but had twice the number of accepted draws.

Responses by variable

Generally, exports react with largest magnitudes to all shocks in the SVAR model, which is in part explained by the high volatility of Finnish foreign trade. Prices have the smallest and in some cases delayed responses, which can be explained in part by frictions. Responses of output and hours are prone to change the direction after their initial reactions to most shocks. The most clearly significant responses are produced by the technology and labour supply shocks, and the most uninformative responses come from the domestic demand shock. All of the responses converge to zero within seven years after the innovation. Some of the responses lay outside of the credible intervals, which is possible because the representative model is not comprised of pointwise median responses, but instead all of the responses come from the same conditionally simulated model chosen by minimizing the cost function in Equation 14.

As interest rate is regarded exogenous in the EMU era, it should not be affected by the shocks. Indeed, interest rate has the smallest responses to all shocks with magnitudes of 0.05–0.25, only one of which is statistically significant. All four shocks have a negligible impact on interest rate also in KOOMA, with magnitudes of 0.01–0.02. As the interest rate was included in SVAR model to serve as an exogenous conditioning variable, it is left out of the further analysis.

More specifically, the impulse responses which are statistically significant on 68 % credibility level for at least four consecutive quarters within the first two years after the innovation, are the responses of output, hours and exports on both supply shocks, responses of prices to technology and export demand shocks, and response of wages to labour supply shock. Responses, which are statistically significant for at least two consecutive quarters, are the responses of output to both demand shocks, response of prices to labour supply shock, response of wages to technology shock, and response of exports to export demand shock. Additionally, responses, which become statistically significant on a delayed manner after the first two years are the responses of output and hours on external demand shock, responses of prices to labour supply shock, and response of wages to technology shock. Table 4 summarizes the variablewise analysis of statistical significance of responses.

	Significant for at least four quarters	Significant for less than a year	Became significant on a delay	Change direction	Not significant	Robustness
Output	T, L, E	D	T	T, L, E		T, (L), D, E
Hours	T, L, E		L, E	T, L	D	T, L, E
Prices	T, L, E	D		T, L		T, L, D, (E)
Wages	T, L		L		D, E	(T)
Interest rate	T		T		L, D, E	
Exports	T, L, E			T, L, E	D	(T), (L), E

Table 4: Variablewise analysis of SVAR responses. T = technology, L = labour supply, D = domestic demand and E = export demand shock. Parentheses indicate that the response is only somewhat robust.

Responses by shocks

Technology shock. A unit technology shock, see Figure 2, results in 0.9 % robust and somewhat persistent lagged increase in output, followed by a somewhat robust decrease of −0.2 % after three years. The two lowermost responses underneath the credible interval come from the models in robustness check 9. After initial and short lived decrease of −0.6 %, hours increase by 0.4 % a year after, both robustly. Prices decrease initially by −0.2 % robustly, and reverse direction after two years, reaching quite persistent increase of 0.1 % with somewhat robustness. The violations in the secondary impact come from robustness checks 7–11. The response of wages diminish slowly after initial increase of 0.5 % losing significance temporarily after the first year, being somewhat robust. Exports increase somewhat robustly by 2 % with a delay after the first year, reversing direction with −0.5 % decrease after three years. Violations come from robustness checks 8–10.

Labour supply shock. Labour supply shock is followed by a somewhat robust increase in output of 0.6 %, with a decline of −0.1 % after three years. Hours similarly rise robustly by 0.3 % with a lag, and fall after three years by −0.1 %, but not robustly. Prices first decline robustly by −0.3 % for two quarters, but change direction later for two years, with the lowermost deviation coming from robustness check 9. Wages have a delayed but not robust increase of 0.2 % for two years. Exports likewise respond on a delay by increasing somewhat robustly by 1.5 %, and declining by −0.2 % after three years.

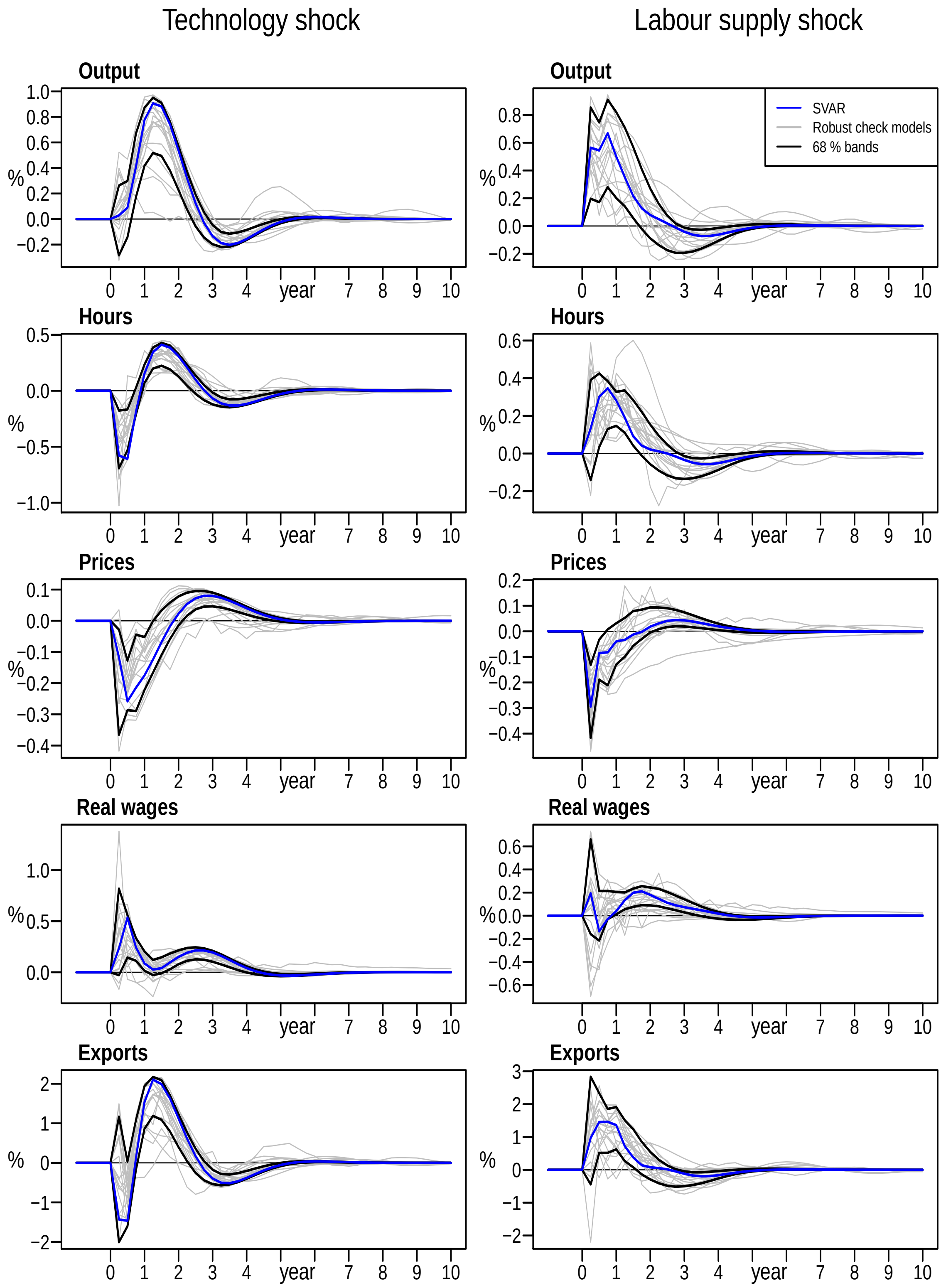

Figure 2: Impulse responses to supply shocks from SVAR -model with 68 % bands.

Domestic demand shock. The response of output to domestic demand shock, see Figure 3, is a sharp and robust increase of 0.4 % for two quarters. Prices rise sharply by 0.4 %, being significant and robust only for the first quarter. The responses of hours, wages and exports are not significant.

Export demand shock. External demand shock results in initial robust 0.8 % positive impact on output, followed by a robust decline of −0.2 % after two years.

Response of hours show lagged robust decline of -0.1 % after two years. Prices increase by 0.1 % for two years in somewhat robust manner. The response of wages is generally not significant, although it does become significantly negative after three years but with negligible magnitude. Exports react with a positive 3 % robust increase in the first year, declining by -0.5 % after two years, upper deviation coming from the robustness check 9.

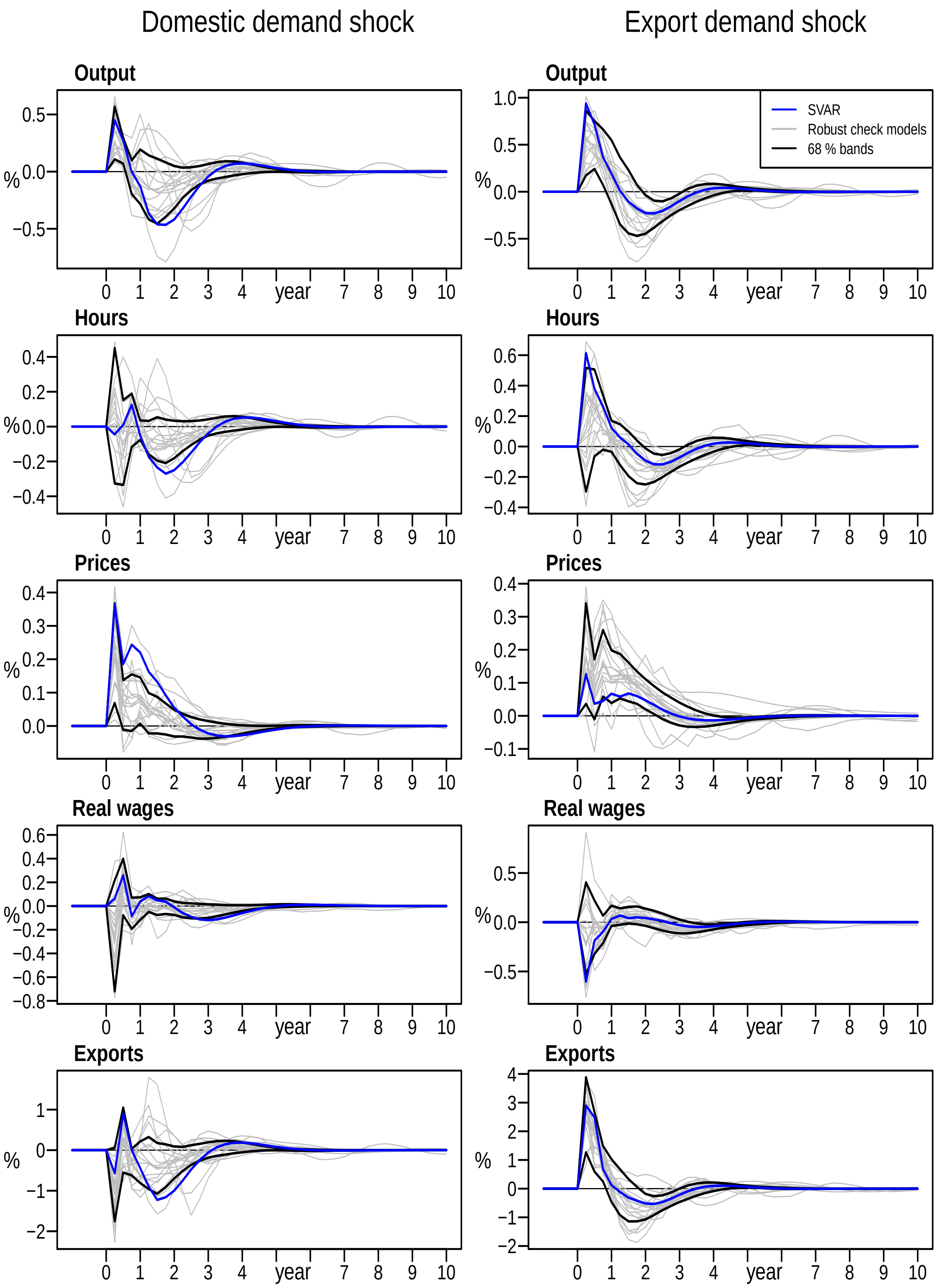

Figure 3: Impulse responses to demand shocks from SVAR -model with 68 % bands.

5 Comparison of KOOMA and SVAR models

In this Section I compare the signs $(+, -)$, magnitudes and persistence of equivalent IRFs in pairs from both models. Also the paths of IRFs are compared in those cases, where the signs of response from either of the models reverses. Impulse responses are regarded similar, if KOOMA's response is included in the credible interval of the SVAR model.

The signs of all pairs of significant responses are the same in both models, except on responses of wages to both supply shocks. The magnitudes of responses to all shocks are generally similar in both models on output, larger in KOOMA with respect to hours, prices and wages, and larger in the SVAR model on exports. The persistence of responses are generally similar on output, hours and exports, and larger in KOOMA on prices and wages.

I will first describe the theoretical economic channels intermediating the impacts of the four shocks in KOOMA one shock at a time, and compare their IRFs to the equivalent IRFs from SVAR model. These economic channels in KOOMA are based on macroeconomic theory and New Keynesian DSGE literature. All responses from KOOMA converge eventually to zero by construction in the General Equilibrium modelling framework. Impulse responses from KOOMA and SVAR models are plotted for 10 year period after the initial innovations to the shock processes. The y-axis values represent approximate percentage changes, as the variables are in logarithmic differences. I have added an **R** or **r** to some of the plots to indicate that their significant SVAR responses are also robust or somewhat robust, respectively. I also labelled those IRFs with **S**, which had sign restrictions issued upon them in the SVAR model.

Technology shock

Technology shock increases marginal productivity of labour and allows firms to meet their supply with less production factors, reducing the need for labour. At the same time firms' marginal costs diminish, increasing aggregate supply and output, and decreasing prices. As the labour productivity increases more than the new optimal supply, the demand for labour input and hours worked drop. As a result wages decline, which has a temporary negative impact on consumption. Aggregate demand and exports increase in response to lowered prices. This, in turn, increases output and labour demand, reversing the sign of hours worked. Lower domestic price level supports exports, output, employment and aggregate demand, and help the economy to converge back to baseline

A unit innovation to technology shock in KOOMA, see Figure 4, increases output for two years, reaching nearly 1 %, before slowly declining back towards zero. The magnitude is exactly the same as in SVAR model, but the persistence is much higher in KOOMA. Hours and prices fall sharply by -1.5 %, former raising back to zero fast and the latter slowly. The path of hours changes signs twice in SVAR model, similarly to KOOMA, but has a smaller scaling factor. Prices has much smaller magnitude and persistence and only changes signs in SVAR model. Wages decline by -2 % with a persistent recovery, being the only somewhat robust response with an opposite sign in SVAR model. It is also accompanied by smaller magnitude and persistence and an unmatched lagged beginning for the response in SVAR model. Exports increase by 1 %, less than in SVAR model, and exert greater persistence. The sign is also only reversed in SVAR model, although the slow increase in KOOMA's response is partially matched by the SVAR response, which becomes significant only after a delay of one year.

Labour supply shock

Labour supply shock increases the willingness of households to supply more labour with all salary levels, leading to reduced wages and firms' marginal costs, and increased employment and output. In consequence prices decrease, increasing aggregate demand and exports similarly as in the case of technology shock.

Unit labour supply shock causes output to increase by 0.2 % in KOOMA, having smaller magnitude but more persistence than in SVAR model. Hours rise immediately by 1 % with persistent recovery in KOOMA, whereas the response only becomes significant after a lag in SVAR model, but magnitudes and paths are similar from thereon. Prices fall by -0.5 % in KOOMA, more than in SVAR model, and exerts higher persistence. Also the sign is only reversed in SVAR model. Wages drop by -1.5 %, being the other response with opposite sign than in SVAR model. Also the magnitudes and persistence are not matched, but this SVAR response was not found robust. The impact on exports is more persistent in KOOMA but with smaller size of 0.3 % than in SVAR model.

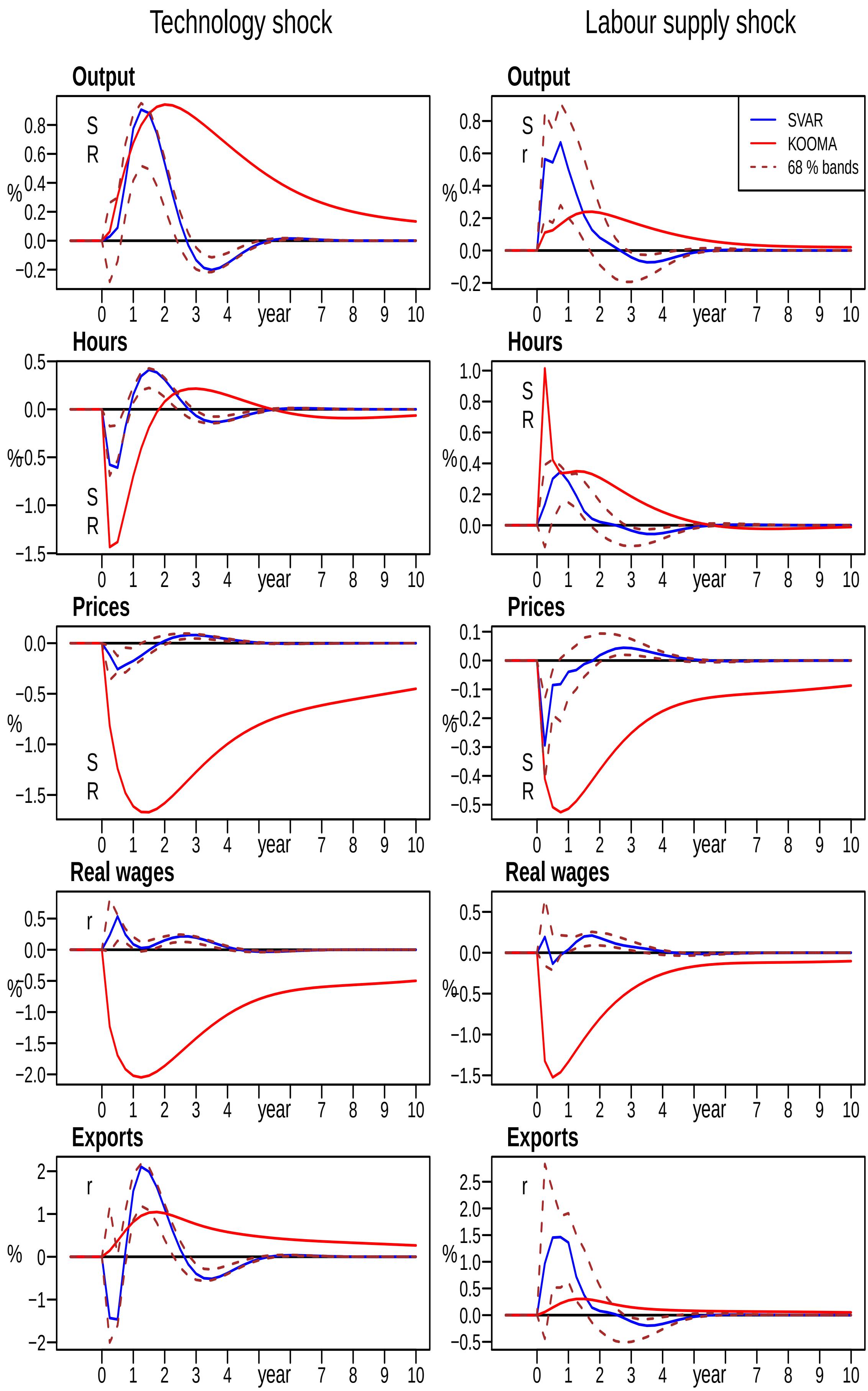

Figure 4: Comparison of supply shocks. I have marked an **S** on the left hand side of those IRFs, which had sign restrictions issued upon them in the SVAR model. The robustness of a significant response is indicated by an upper case **R**, while a lower case **r** denotes for a somewhat robust response.

Domestic demand shock

Domestic demand shock increases the marginal utility of consumption and the marginal rate of substitution between consumption and leisure. The change in marginal rate of substitution causes wages to increase, which is passed on to increased prices. Higher wages also increase firms' marginal costs. The elevated consumption increases output and offsets the decline in exports, which is caused by the weakened cost competitiveness.

The initial response of output to domestic demand shock, Figure 5, has the same magnitude of 0.4 % in both models, but has more persistence in KOOMA, and the SVAR model fails to produce similar path as the reversed sign has no significant significance. Hours increase by 0.5 % in KOOMA and reverses sign after two years. The SVAR response is not significant and the response from KOOMA exits the credible bands after the initial reaction. Prices respond with equal magnitude of 0.4 % in both models but the response of KOOMA build up slower and exerts more persistence. Response of wages increase slowly in KOOMA and reach 0.6 %, being outside of the bands, although SVAR response is not significant. In KOOMA, response of exports decline slowly by -0.2 % with a persistent recovery, but is contained within the SVAR bands for the first three years. The SVAR response is not significant here either.

Export demand shock

In KOOMA, External demand shock increases the exports directly and thereby also output increases immediately. The increased foreign demand creates wage pressures, as the demand for labour rises. Consumer price level will rise because of increased demand but less than nominal wages, and hence real wages will rise. This has positive impact for consumption. Rise in domestic prices reduces the cost competitiveness, and weakens the increased demand.

External demand shock increases all variables in KOOMA. Output increases by 1 %, having nearly identical magnitudes and paths with the response from SVAR model, and almost the same persistence. Hours rise by 0.9 % in KOOMA and reverses sign to -0.2 % on the second year. The initial response is larger than the non-significant response in SVAR model, but the secondary impact has similar sizes in both models, but more persistence in KOOMA. Prices rise by 0.5 %, much more than the SVAR response suggest. Also the persistence is higher in KOOMA, although the SVAR response is unusually persistent here. Wages rise by 1 % in KOOMA, with persistent decline, whereas SVAR response is smaller and insignificant. Exports behave quite similarly in both models, increasing briefly by 4 % in KOOMA having a bit higher magnitude and persistence than SVAR model, and reversing signs with smaller magnitude in KOOMA.

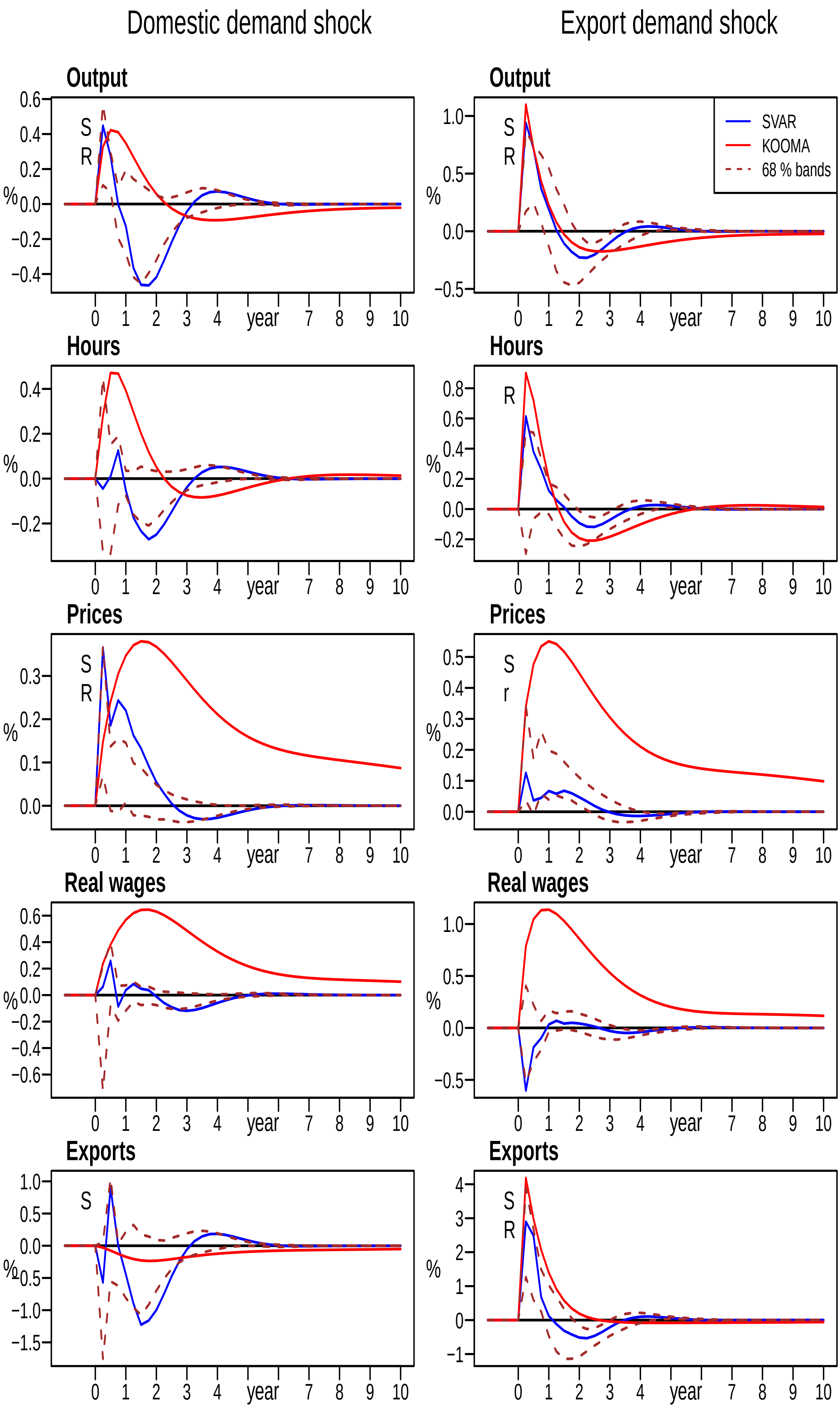

Figure 5: Comparison of demand shocks.

6 Conclusions

The aim of this study is to evaluate the calibration of the large macro model KOOMA of the Ministry of Finance of Finland. KOOMA is an attempt to model the Finnish economy in a popular Dynamic Stochastic General Equilibrium (DSGE) framework. In order to be a useful tool in forecasting and policy analysis, the model should be able to emulate how certain important macroeconomic phenomena behave in Finland. I evaluate the model by comparing chosen impulse response functions from KOOMA to equivalent significant and robust impulse responses obtained from a data driven Structural VAR model. The significance of the responses within the SVAR model, and the significance of the differences between responses from KOOMA and SVAR models, are determined on the basis of one standard deviation credible intervals with a Bayesian interpretation.

I deliberately take several steps along this study to reduce the risk of looking only for results which conform to a priori expectations: choosing only two lags for the VAR model to reduce the number of estimated parameters. Using sign restrictions to identify the SVAR model, with sum specification concerning only the net impact of shocks over an interval of time, both of which allow more freedom for the data to manifest. The use of only one standard deviation as the width of the credible intervals, restraining from concluding anything about the tail events, or expressing unreasonably high confidence in the results. Conducting extensive robustness analysis to find out how consistent the SVAR responses remain when the variable and model specifications are altered in various ways.

The results indicate, that KOOMA generally produces impulse responses with the same signs as the SVAR model. Output responds with similar magnitudes to shocks in both models, while KOOMA produces larger magnitudes for responses of hours, prices and wages, and smaller magnitudes for responses of exports. The persistence of responses are generally similar on output, hours and exports, and larger in KOOMA on prices and wages.

However, because of sampling and model related uncertainty and volatility in the data, the results are not fully conclusive. The questions of whether some parts of KOOMA should be re-calibrated and in what way are a topic for another study.

References

Ahola, I. (2012). Suomen suhdannevaihteluiden tyylitellyt faktat. Master's thesis. University of Helsinki.

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews* **26**(2-4), 113–172.

Baumeister, C. and Hamilton, J. D. (2015). Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *Econometrica* **83**(5), 1963–1999.

Bernanke, B. S., Boivin, J. and Eliasz, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *The Quarterly Journal of Economics* **120**(1), 387–422.

Blanchard, O. J. (2018). On the future of macroeconomic models. *Oxford Review of Economic Policy* **34**(1-2), 43–54.

Blanchard, O. J. and Quah, D. (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review* **79**(4), 655–673.

Blanchard, O. and Perotti, R. (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *The Quarterly Journal of Economics* **117**(4), 1329–1368.

Brüggemann, R. and Kascha, C. (2017). Directed Graphs and Variable Selection in Large Vector Autoregressive Models. Working Paper Series of the Department of Economics, University of Konstanz 2017-06. Department of Economics, University of Konstanz.

Canova, F., Lopez-Salido, D. and Michelacci, C. (2013). The Ins and Outs of Unemployment: An Analysis Conditional on Technology Shocks. *Economic Journal* **123**, 515–539.

Canova, F. and Nicolo, G. D. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics* **49**(6), 1131–1159.

Canova, F. and Paustian, M. (2011). Business cycle measurement with some theory. *Journal of Monetary Economics* **58**(4), 345–361.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy* **113**(1), 1–45.

Christiano, L. J., Eichenbaum, M. and Trabandt, M. (2018). On DSGE Models. *Journal of Economic Perspectives* **32**(3), 113–140.

Christiano, L. J., Eichenbaum, M. and Vigfusson, R. (2007). Assessing Structural VARs. *in* NBER Macroeconomics Annual 2006, Volume 21. NBER Chapters. National Bureau of Economic Research, Inc. pp. 1–106.

Consolo, A., Favero, C. A. and Paccagnini, A. (2009). On the statistical identification of DSGE models. *Journal of Econometrics* **150**(1), 99–115.

Del Negro, M. and Schorfheide, F. (2004). A DSGE-VAR for the Euro Area. Computing in Economics and Finance 2004 79. Society for Computational Economics.

Del Negro, M. and Schorfheide, F. (2006). How good is what you've got? DSGE-VAR as a toolkit for evaluating DSGE models. *Economic Review* (Q 2), 21–37.

Del Negro, M., Schorfheide, F., Smets, F. and Wouters, R. (2007). On the Fit of New Keynesian Models. *Journal of Business & Economic Statistics* **25**, 123–143.

Doan, T., Litterman, R. and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* **3**(1), 1–100.

Faust, J. (1998). The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy* **49**(1), 207–244.

Fernández-Villaverde, J. and Rubio-Ramírez, J. F. (2008). How Structural Are Structural Parameters?. *in* NBER Macroeconomics Annual 2007, Volume 22. NBER Chapters. National Bureau of Economic Research, Inc. pp. 83–137.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J. and Watson, M. W. (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review* **97**(3), 1021–1026.

Fernández-Villaverde, J., Rubio-Ramírez, J. F. and Schorfheide, F. (2016). *Solution and Estimation Methods for DSGE Models*. Vol. 2 of *Handbook of Macroeconomics*. Elsevier. chapter 0, pp. 527–724.

Fry, R. and Pagan, A. (2011). Sign Restrictions in Structural Vector Autoregressions: A Critical Review. *Journal of Economic Literature* **49**(4), 938–960.

Fève, P., Matheron, J. and Sahuc, J.-G. (2010). Disinflation Shocks in the Eurozone: A DSGE Perspective. *Journal of Money, Credit and Banking* **42**(2-3), 289–323.

Gali, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?. *American Economic Review* **89**(1), 249–271.

Granziera, E., Moon, H. R. and Schorfheide, F. (2018). Inference for VARs identified with sign restrictions. *Quantitative Economics* **9**(3), 1087–1121.

Gulan, A., Haavio, M. and Kilponen, J. (2014). Kiss me deadly: From Finnish great depression to great recession. Research Discussion Papers 24/2014. Bank of Finland.

Hendry, D. F. and Muellbauer, J. N. J. (2018). The future of macroeconomics: macro theory and models at the Bank of England. *Oxford Review of Economic Policy* **34**(1-2), 287–328.

Kilian, L. and Lutkepohl, H. (2017). *Structural Vector Autoregressive analysis (Themes in Modern Econometrics)*. Cambridge: Cambridge University Press.

Kimball, M. S., Fernald, J. G. and Basu, S. (2006). Are Technology Improvements Contractionary?. *American Economic Review* **96**(5), 1418–1448.

Kydland, F. E. and Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica* **50**(6), 1345–1370.

Lanne, M. and Lütkepohl, H. (2008). Identifying Monetary Policy Shocks via Changes in Volatility. *Journal of Money, Credit and Banking* **40**(6), 1131–1149.

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control* **34**(2), 121–131.

Leeper, E. M., Walker, T. B. and Yang, S. S. (2013). Fiscal Foresight and Information Flows. *Econometrica* **81**(3), 1115–1145.

Lehmus, M. (2014). Finnish fiscal multipliers with a structural VAR model.

Levintal, O. (2017). Fifth-order perturbation solution to DSGE models. *Journal of Economic Dynamics and Control* **80**(C), 1–16.

Lucas Jr., R. E. and Sargent, T. J. (1979). After Keynesian macroeconomics. *Quarterly Review* (Spr).

Lucas, R. J. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* **1**(1), 19–46.

Mandelbrot, B. (2008). *Misbehavior of Markets: A Fractal View f Risk, Ruin and Reward*. Profile Books Ltd.

Pagan, A. and Robinson, T. (2016). Investigating the Relationship Between DSGE and SVAR Models. NCER Working Paper Series 112. National Centre for Econometric Research.

Peersman, G. and Straub, R. (2006). Putting the New Keynesian Model to a Test. Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 06/375.

Peersman, G. and Straub, R. (2009). Technology Shocks And Robust Sign Restrictions In A Euro Area Svar. *International Economic Review* **50**(3), 727–750.

Ravenna, F. (2007). Vector autoregressions and reduced form representations of DSGE models. *Journal of Monetary Economics* **54**(7), 2048–2064.

Rigobon, R. (2003). Identification Through Heteroskedasticity. *The Review of Economics and Statistics* **85**(4), 777–792.

Romer, P. (2016). The Trouble with Macroeconomics. Commons memorial lecture of the omicron delta epsilon society.

Rubio-Ramírez, J. F., Waggoner, D. F. and Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *Review of Economic Studies* **77**(2), 665–696.

Sariola, M. (2015). What drives business cycles in Sweden? A sign restriction structural VAR approach. Technical report. Discussion Papers of the Ministry of Finance of Finland.

Sax, C. and Eddelbuettel, D. (2018). Seasonal adjustment by x-13arima-seats in r. *Journal of Statistical Software, Articles* **87**(11), 1–17.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica* **48**(1), 1–48.

Sims, C. A. (2012*a*). Statistical Modeling of Monetary Policy and Its Effects. *American Economic Review* **102**(4), 1187–1205.

Sims, C. A. and Zha, T. (2006). Were There Regime Switches in U.S. Monetary Policy?. *American Economic Review* **96**(1), 54–81.

Sims, E. R. (2012*b*). News, Non-Invertibility, and Structural VARs. Working Papers 013. University of Notre Dame, Department of Economics.

Smets, F. and Wouters, R. (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association* **1**(5), 1123–1175.

Solow, R. (2010). Building a Science of Economics for the Real World. Congressional testimony.

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy* **34**(1-2), 70–106.

Taleb, N. (2007). *Black Swan: The Impact of the Highly Improbable*. Random House. New York.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics* **52**(2), 381–419.

Vaughn, K. (2013). Hayek, Equilibrium, and The Role of Institutions in Economic Order. *Critical Review* **25 (3-4)**, 473–496.

Appendix

Here I detail out the 15 robustness check models, and plot the results from the individual comparisons to the representative model. I also plot all of the comparisons between KOOMA and SVAR models, the SVAR responses with 90 % credible intervals and IRFs from the two unidentified structural shocks in the model.

Altering the SVAR model in robust checks 1–6

Robust check 1, three best models. The first comparison, see Figure 6, is between the three models with smallest summed squared distances to median responses in each IRF among the original 5000 simulated accepted models. Responses from both secondary models are contained in the credible bands of the representative model in almost every IRF.

Robust check 2, sign restrictions only on the first period. Here I test the sensitivity of the responses to the chosen sign restriction scheme, see Figure 7, i.e. the requirement that the net impact of constrained IRFs should be of chosen sign on the four period interval. The alternative scheme is to only restrict the signs of responses on the first period. The responses are otherwise similar in both specifications, but the response of prices to export demand shock seems to be less persistent in the alternative model.

Robust check 3, sign restrictions on wages instead of hours. In this alternative structural model, Figure 8, I have identified the supply shocks by issuing sign restrictions on wages instead of hours, as in Peersman and Straub (2009). Wages are assumed to increase after the technology shock and to diminish after the labour supply shock. This does not cause changes in the responses of hours to supply shocks or in the response of wages to technology shock, but the initial response of wages to labour supply shock is significantly lower.

Robust check 4, posterior medians. Here I compare the SVAR responses to median responses from each period among the 5000 simulated models, Figure 9. By definition, the posterior medians are always contained within the credible intervals. Large differences between the representative responses and medians would indicate a bias in the model (Fry and Pagan, 2011). In most of the cases, the medians are equal or non-significantly smaller in magnitude than the representative responses.

Robust check 5, different number of accepted simulations. The procedure is otherwise unaltered, but I require the simulation algorithm to find 500 and 2500 accepted structural models instead of 5000, Figure 10. The responses remain similar regardless of the number of draws. It is noteworthy, that the 68 % bands obtained from posterior distributions, approximated by different number of draws, are almost identical. Also the 90 % bands are highly similar. This suggests, that the posterior distributions converge quickly with quite low number of draws, at least with these specifications.

Robust check 6, different seed. Similarly to the previous robustness check, Figure 11, I analyze here how much the SVAR responses vary due to factors which are not related to data, by changing the seed of the random number generator of the simulation algorithm from 1917 to 1809. Changing the seed doesn't cause changes in the significant responses.

Altering the VAR model in robust checks 7–9

Robust check 7, different number of lags. In this robust check, Figure 12, I vary the number of lags in estimating the VAR model, using three and four lags instead of two. The significant responses still remain significant, except for the responses of prices to technology and export demand shocks in the case of four lag model. IRFs from both alternate models are more volatile, and the responses from model with four lags doesn't converge to zero as fast as in the other two models. I suspect one reason for the behaviour of these models could derive from the time series, which still might not be completely free of seasonal components. Increasing the number of parameters by adding lags might allow for the models a chance to overfit the in sample, deploying also imaginary roots to match the non-stochastic seasonal behaviour in data.

AIC=6, HQ=1, SC=1, FPE=6

	1	2	3	4	5	6
AIC(n)	−54.39	−54.61	−54.56	−55.28	−55.83	−56.53
HQ(n)	−53.84	−53.59	−53.06	−53.32	−53.40	−53.62
SC(n)	−53.00	−52.02	−50.78	−50.31	−49.66	−49.16
FPE(n)	0.00	0.00	0.00	0.00	0.00	0.00

Table 5: Suggestions for VAR lag count from the range of one to six lags, from four information criteria: Akaike (AIC), Hannah-Quinn (HQ), Schwartz/Bayesian (SC) and Forecast prediction error (FPE). Values of FPE are in the magnitude of e^{-24}.

Robust check 8, data used from 1990 instead of 1999. Here I use the whole available data, Figure 13, starting from the first quarter of 1990. The responses are otherwise similar, but the initial responses of hours and wages to technology shock is larger in the alternate model, and smaller with regards to lagged responses of prices and exports to technology shock. The impacts of labour supply shock are generally more persistent, and also larger on hours.

Robust check 9, pre and post crises data. The data is divided into pre and post financial crises periods, 1999 Q1 – 2008 Q2 and 2008 Q3 – 2017 Q4, respectively. I then estimate VAR models and identify SVAR models separately from both subsets of data, see Figure 14. The resulting models have differing behaviour especially with regards to supply shocks. Both models have smaller magnitudes on responses of output, wages and exports to technology and labour supply shocks and larger magnitudes on hours and prices to external demand shock. They also show higher persistence on responses of hours and prices to labour supply and export demand shocks. The model based on only pre-crisis data produces more volatile responses to labour supply and domestic demand shocks, hinting at the presence of an imaginary root, with similar remarks as with regards to the behaviour of four lag model in Robust check 7.

Altering the variable selection and transformations in robust checks 10–15

Robust check 10, trend gap instead of Year on Year differences. Here I specify the variable transformation as deviations from the Hodrick-Prescott trend path, constructed with lambda of 1600, see Figure 15. Otherwise the responses are similar, except all variables respond to technology shock with smaller magnitudes in this alternative model.

Robust check 11, different measure for inflation. Instead of private consumption prices, I proxy the inflation by the general consumption prices, see Figure 16. Labour supply shock seems to have more persistence and export demand shock larger magnitudes in the alternative model.

Robust check 12, employment instead of hours. I have replaced the total hours worked here with employment, Figure 17. Changes in employment are typically accompanied by changes in total hours worked, but the variables can also change independently, depending on the type of the shock affecting labour markets. The results are similar, but the response of prices to labour supply shock don't seem to reverse signs in the altered model.

Robust check 13, hours as smoothed by StatFin. In this model, Figure 18, I use the total hours worked series which is seasonally adjusted by Statistics Finland, instead of the series which I adjusted myself according to Ahola (2012). The response of prices to domestic demand shock, which is only significant for one quarter in the original model, has smaller magnitude in the altered model. Otherwise the responses are quite similar, although less smooth in the model in the alternative model.

Robust check 14, external demand instead of exports. I replace exports by external demand, Figure 19, which refers to the combined imports of all countries. Although these variables correlate, Finnish exports are also influenced for example by issues of domestic competitiveness, whereas external demand is regarded as completely exogenous for a small open economy. Results remain similar, with responses of exports to supply shocks having a bit less persistence in the altered model.

*Robust check 15, real interest rate *100.* I multiple real interest rate by 100 to see the impact of not having scaled it similarly to real wages, Figure 20. The main difference of this specification is the anticipated increases in magnitudes on responses of interest rate to all shocks. Also the initial response of exports to labour supply shock is significantly negative, whereas the original response only reacts after a lag.

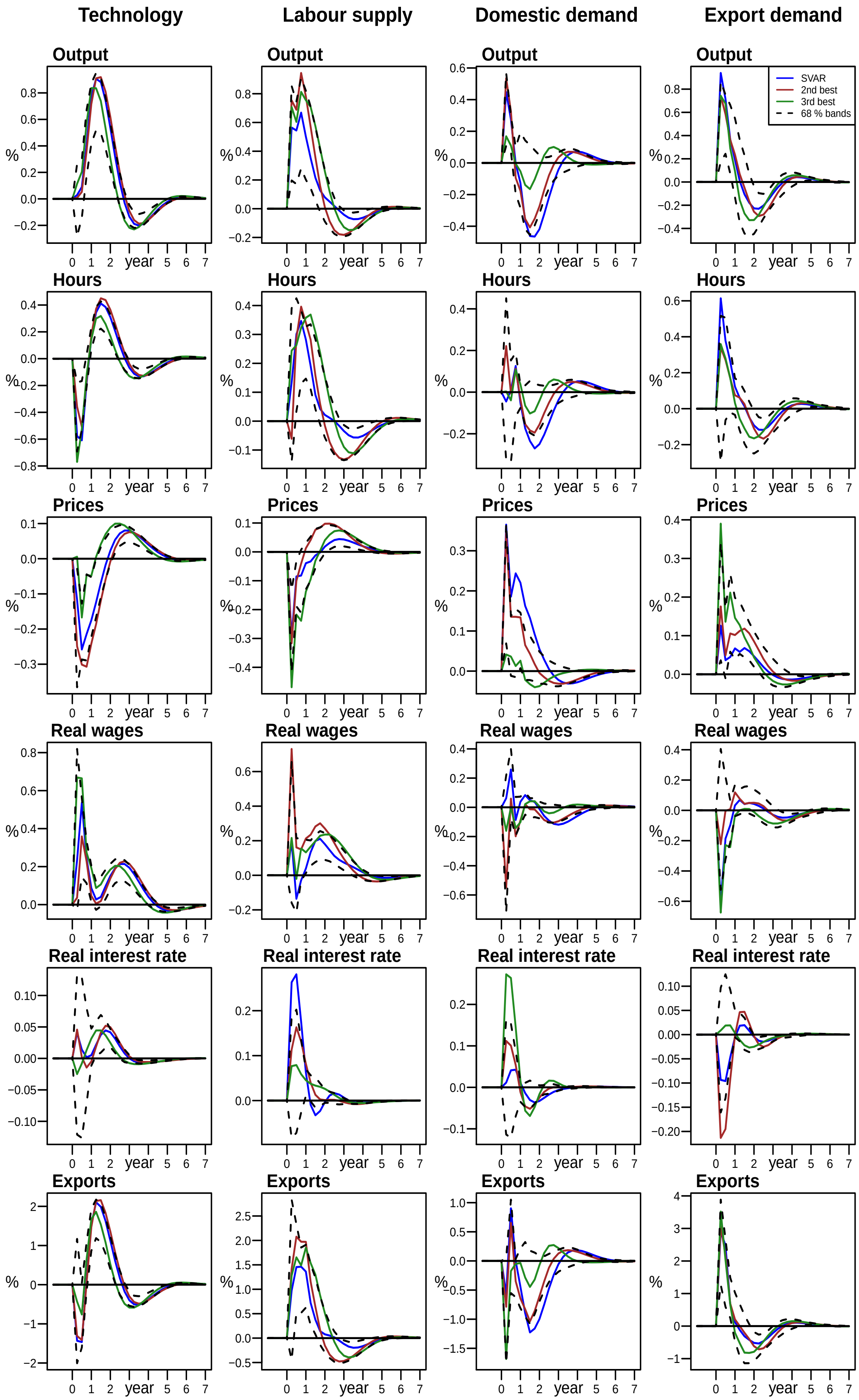

Figure 6: Robust check 1, three best models

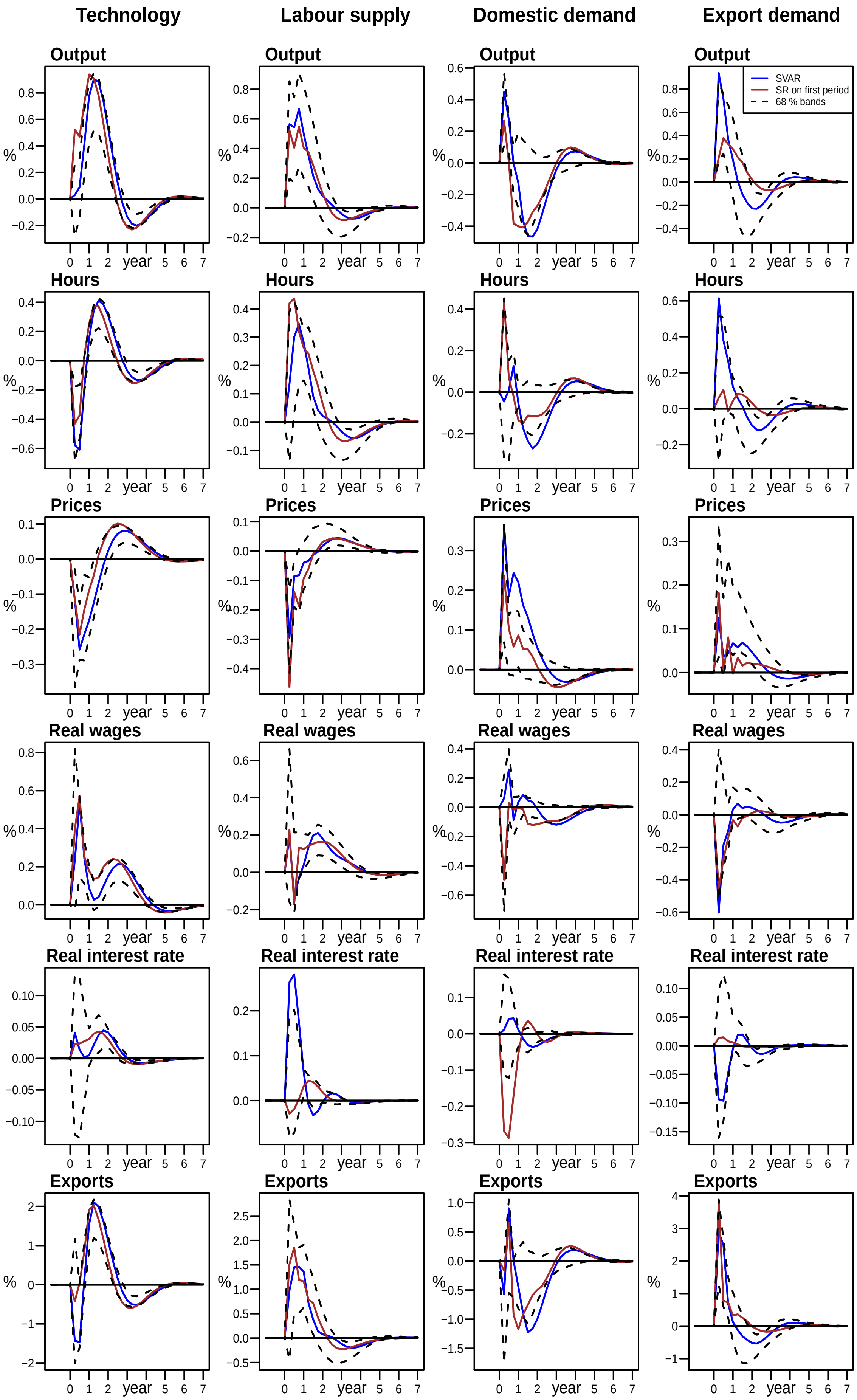

Figure 7: Robust check 2, sign restrictions apply only for the first period.

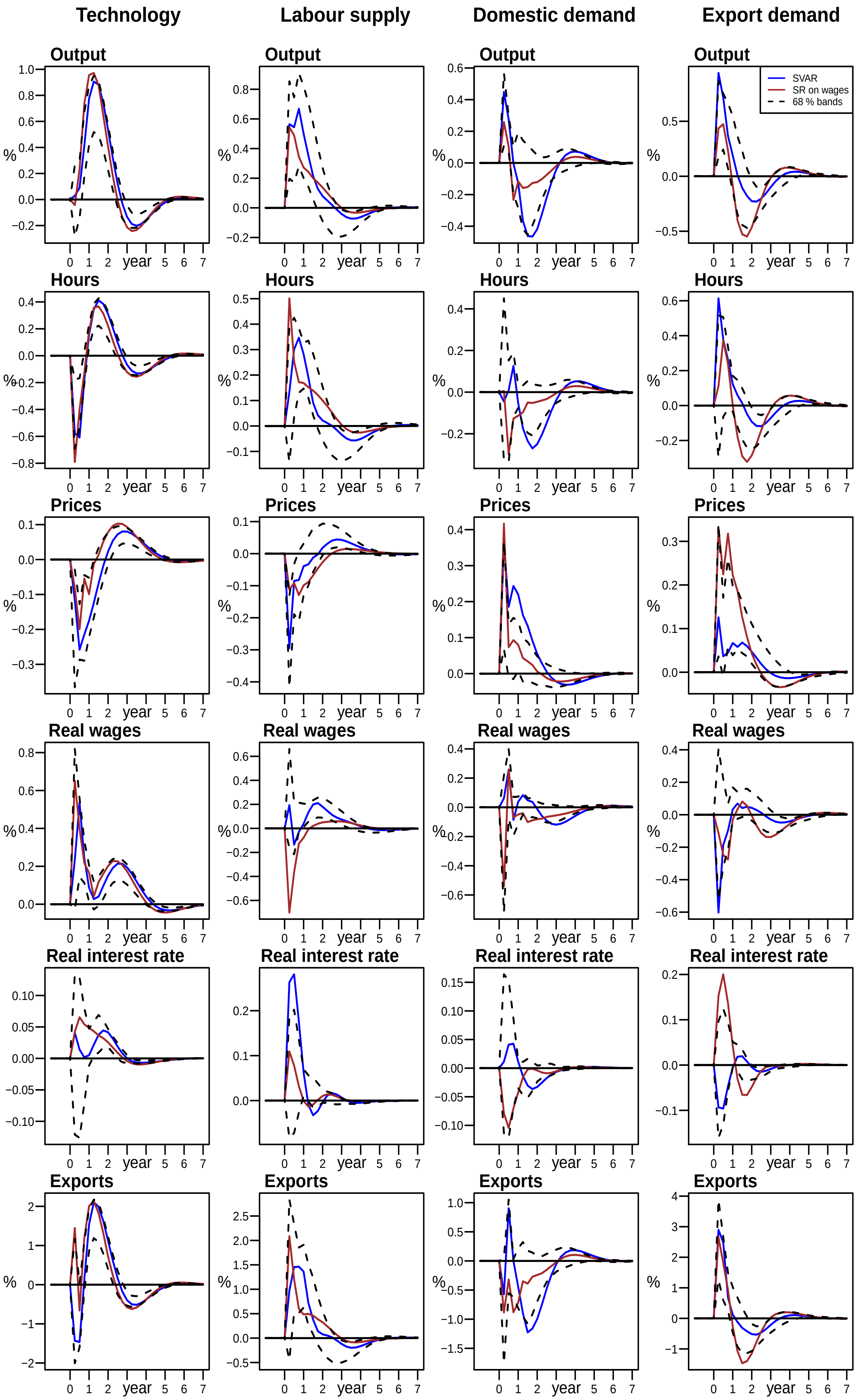

Figure 8: Robust check 3, sign restrictions on wages instead of hours.

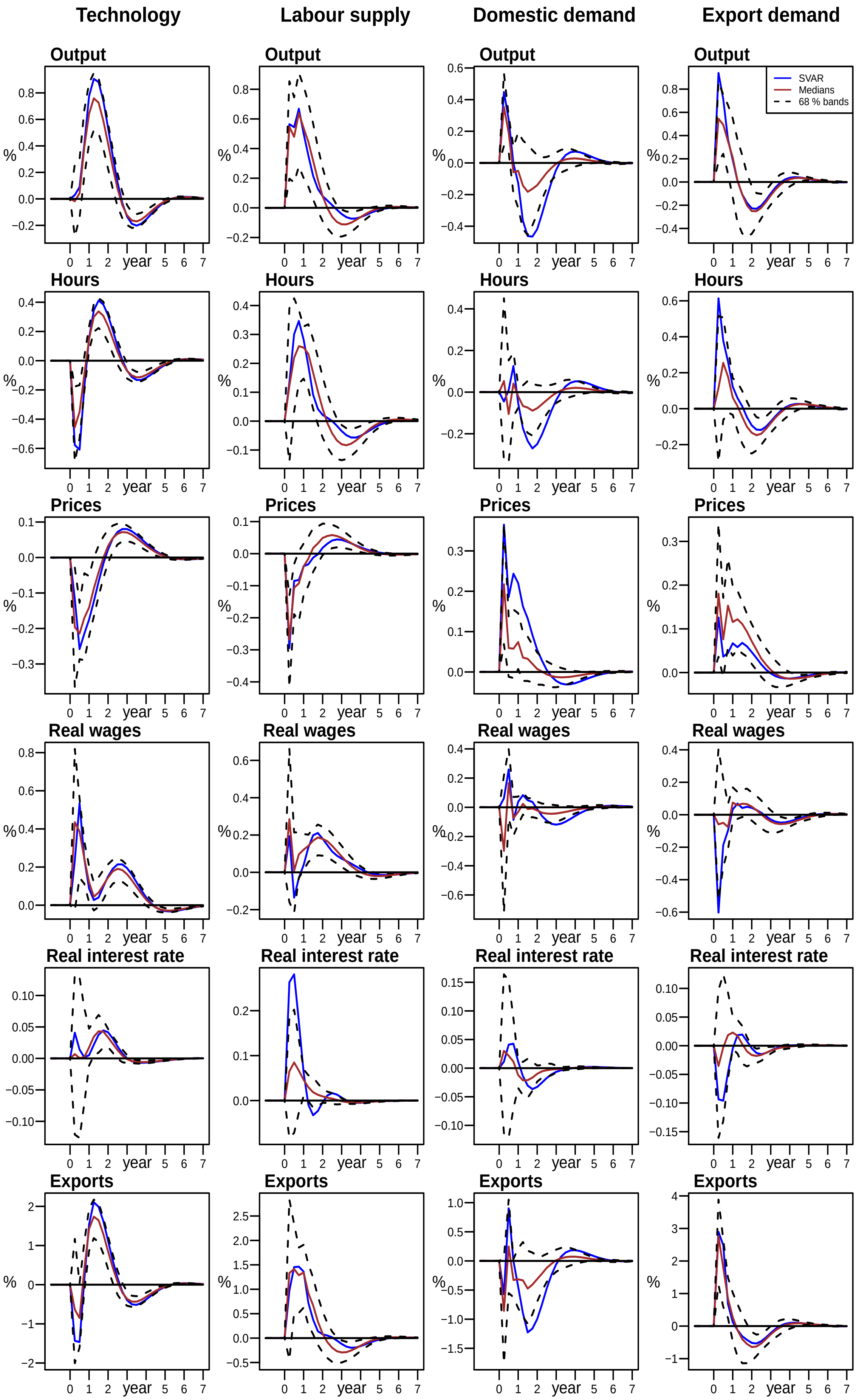

Figure 9: Robust check 4, posterior medians.

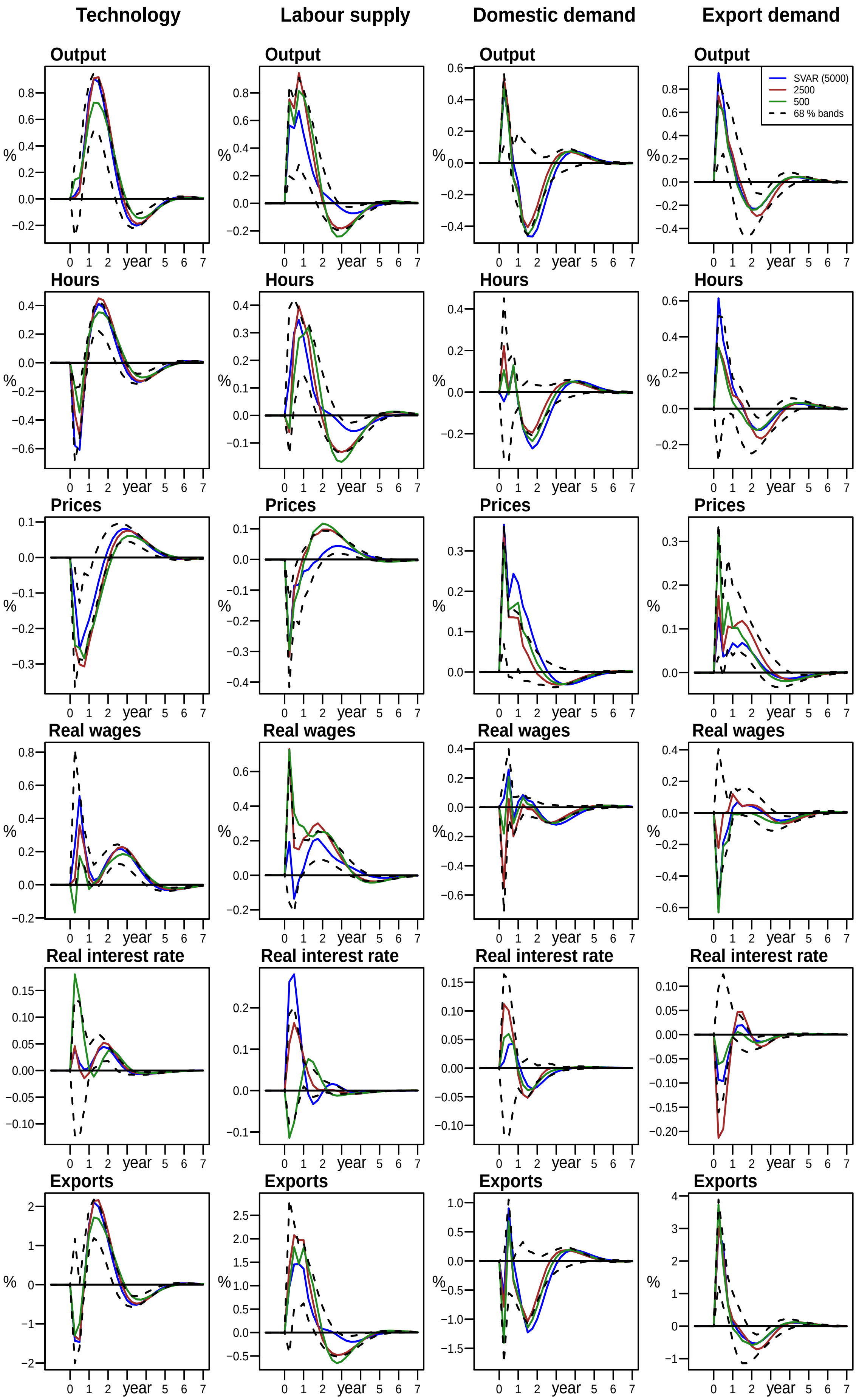

Figure 10: Robust check 5, different required number of accepted simulations.

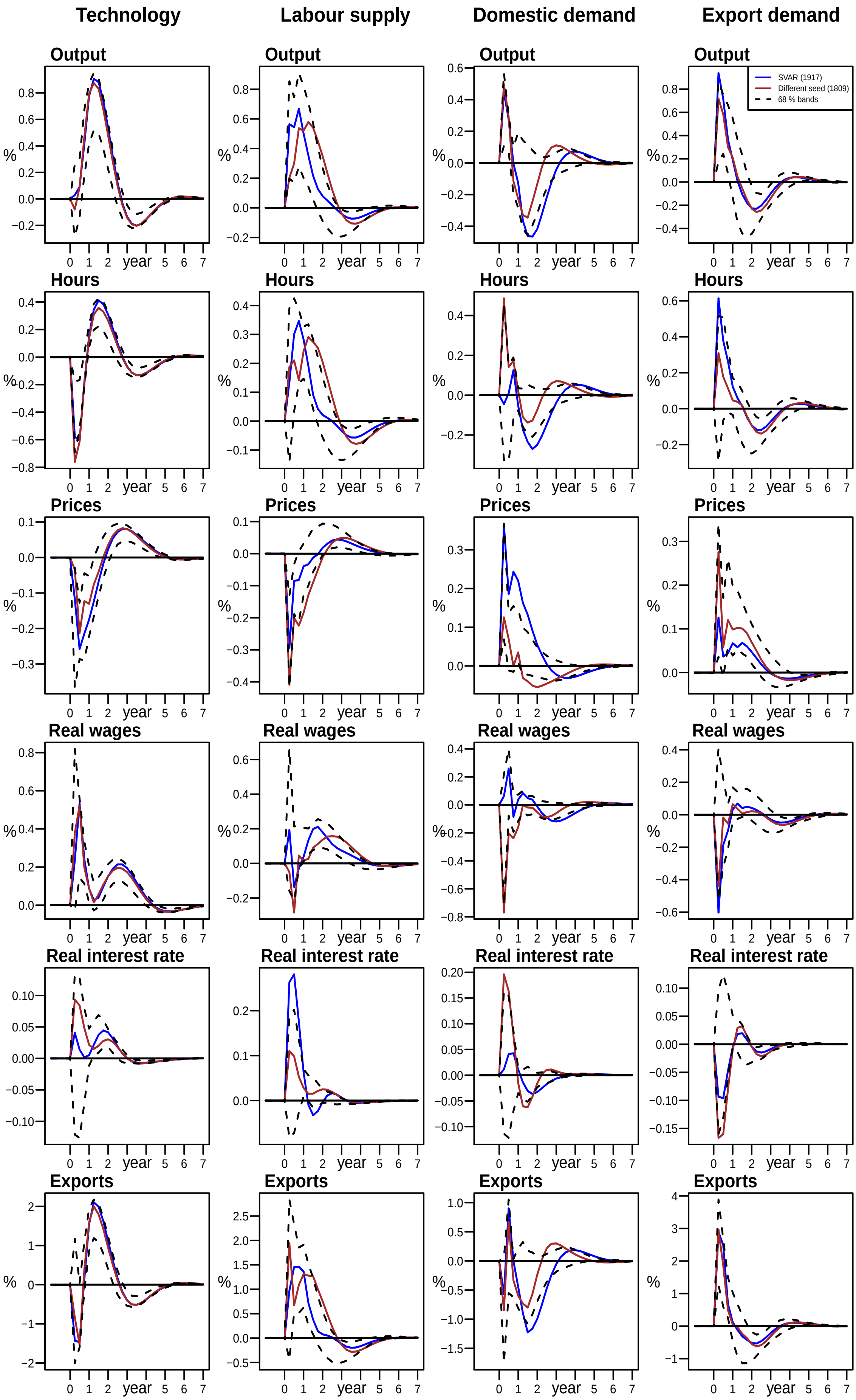

Figure 11: Robust check 6, different seed.

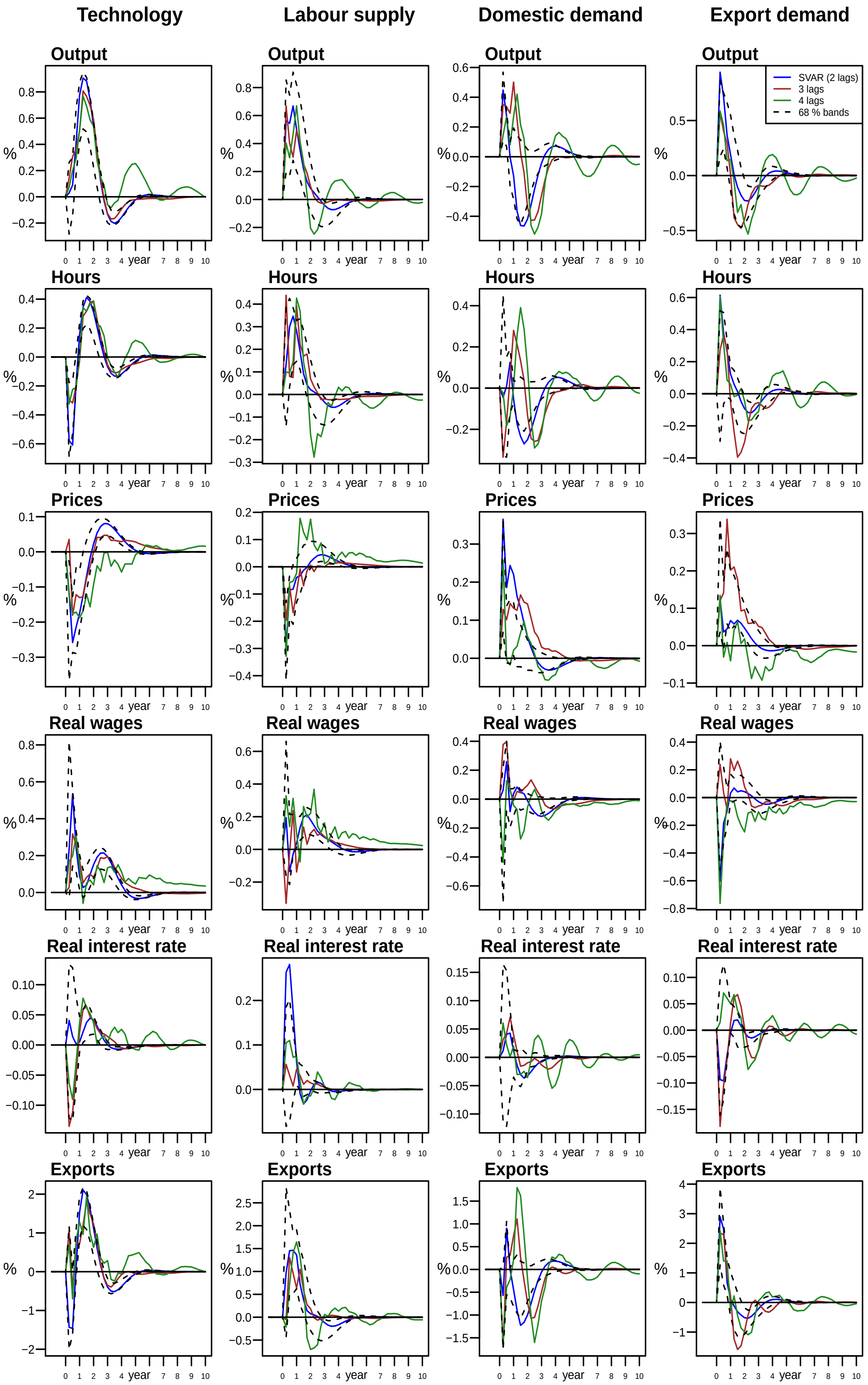

Figure 12: Robust check 7, different number of lags. These plots has a longer time window than other robustness check plots, because the alternative models have more persistent responses.

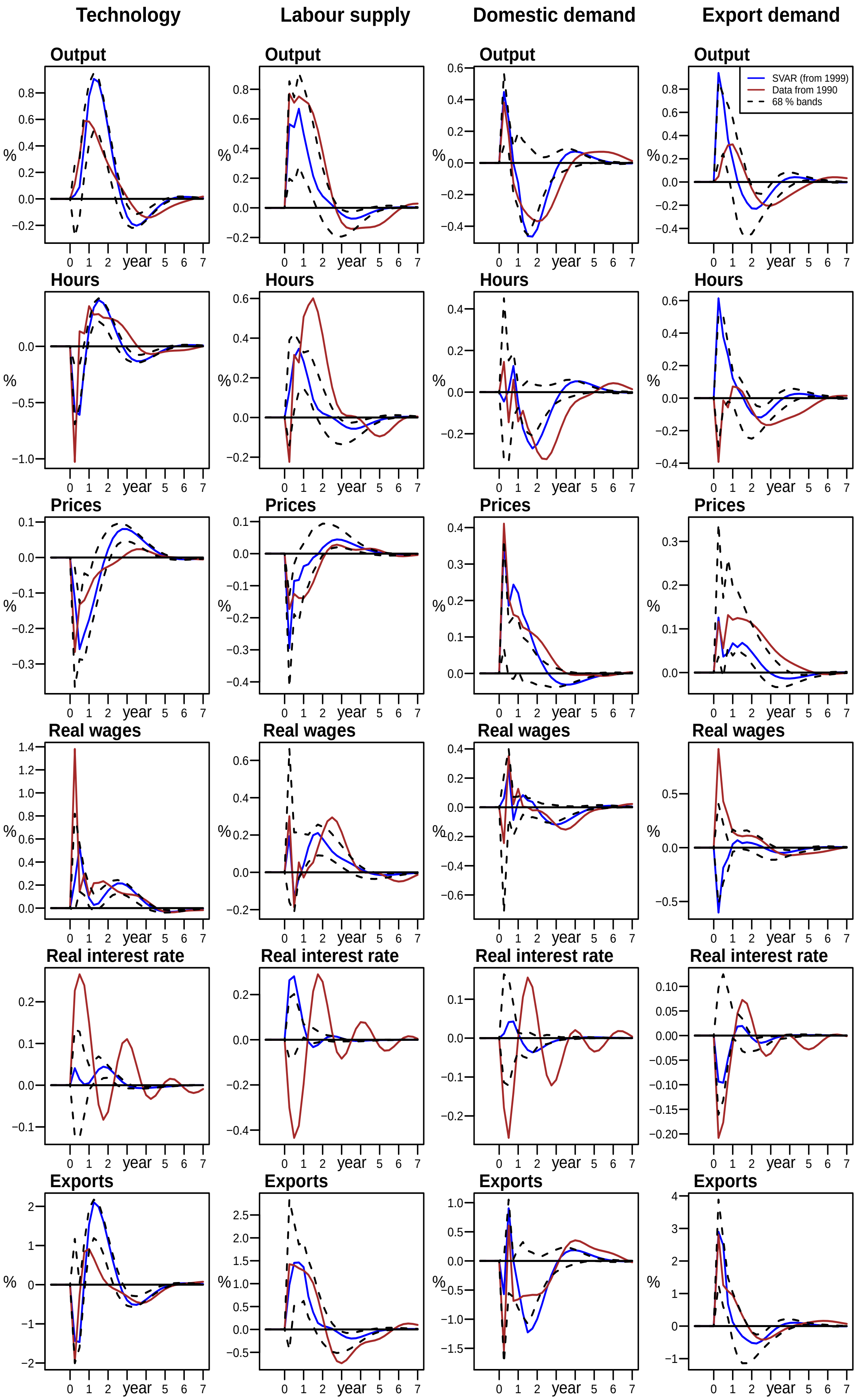

Figure 13: Robust check 8, data used from 1990 instead of 1999.

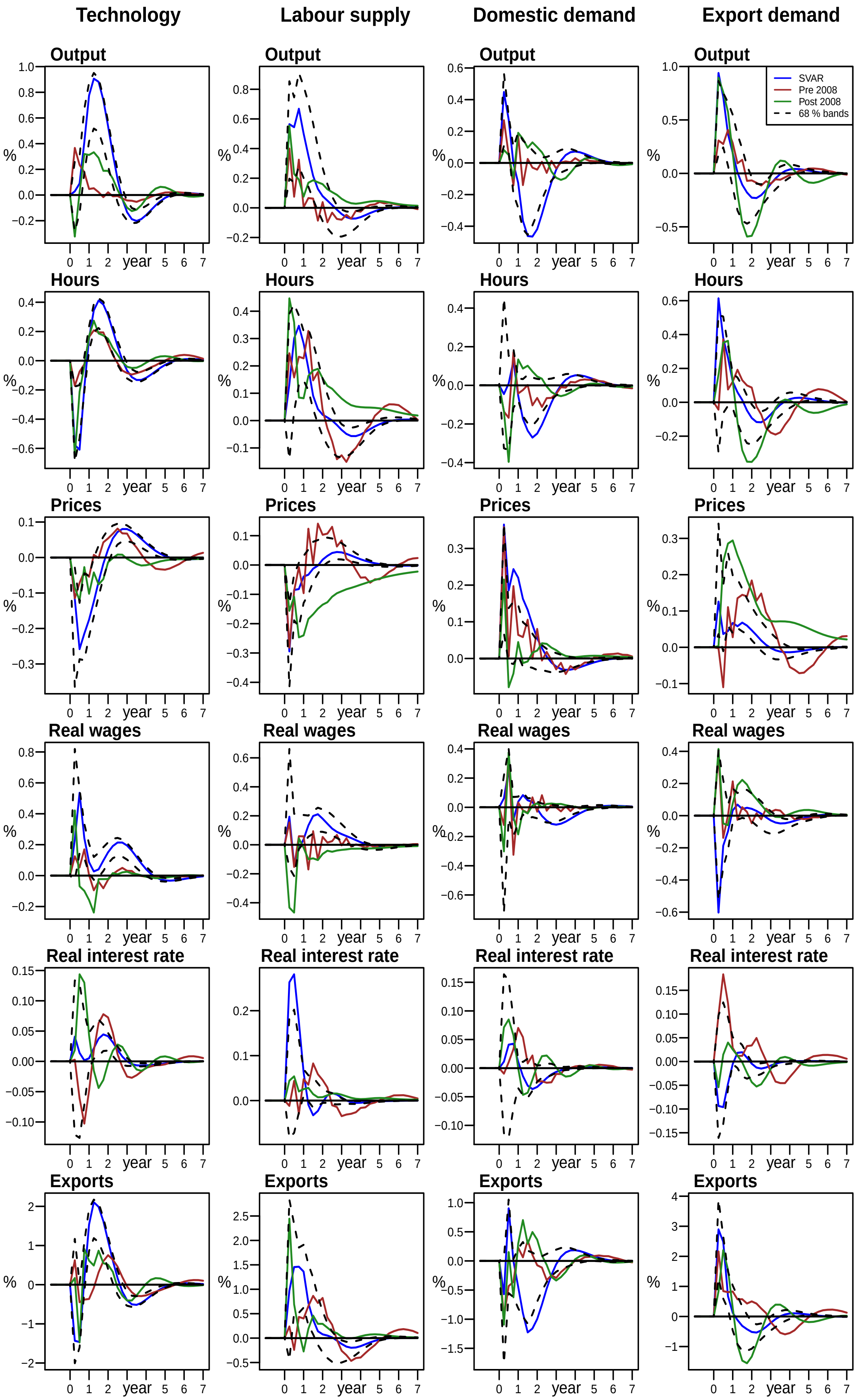

Figure 14: Robust check 9, pre and post crisis data used separately.

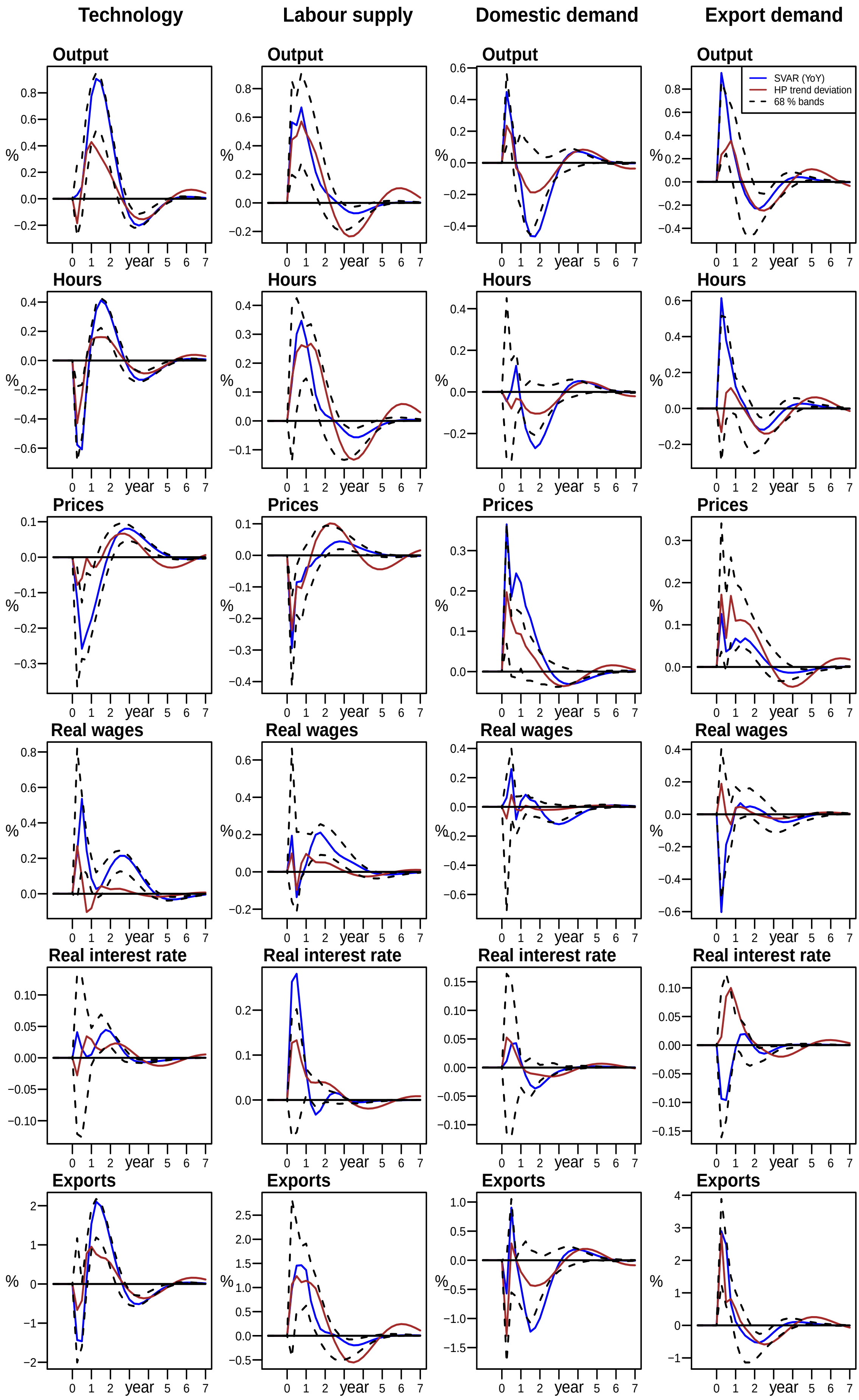

Figure 15: Robust check 10, variables in trend gaps instead of Year on Year differences.

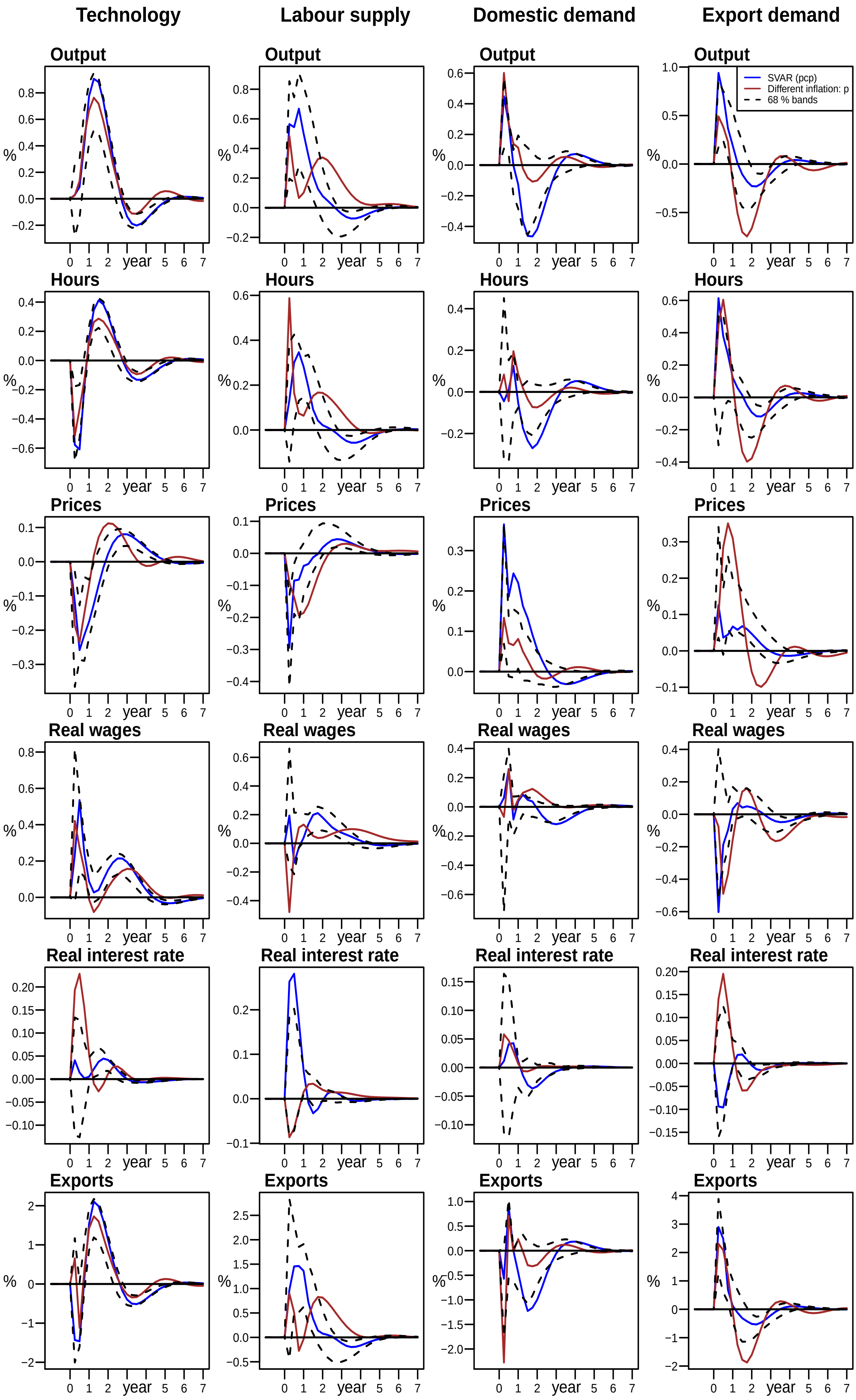

Figure 16: Robust check 11, different measure for inflation.

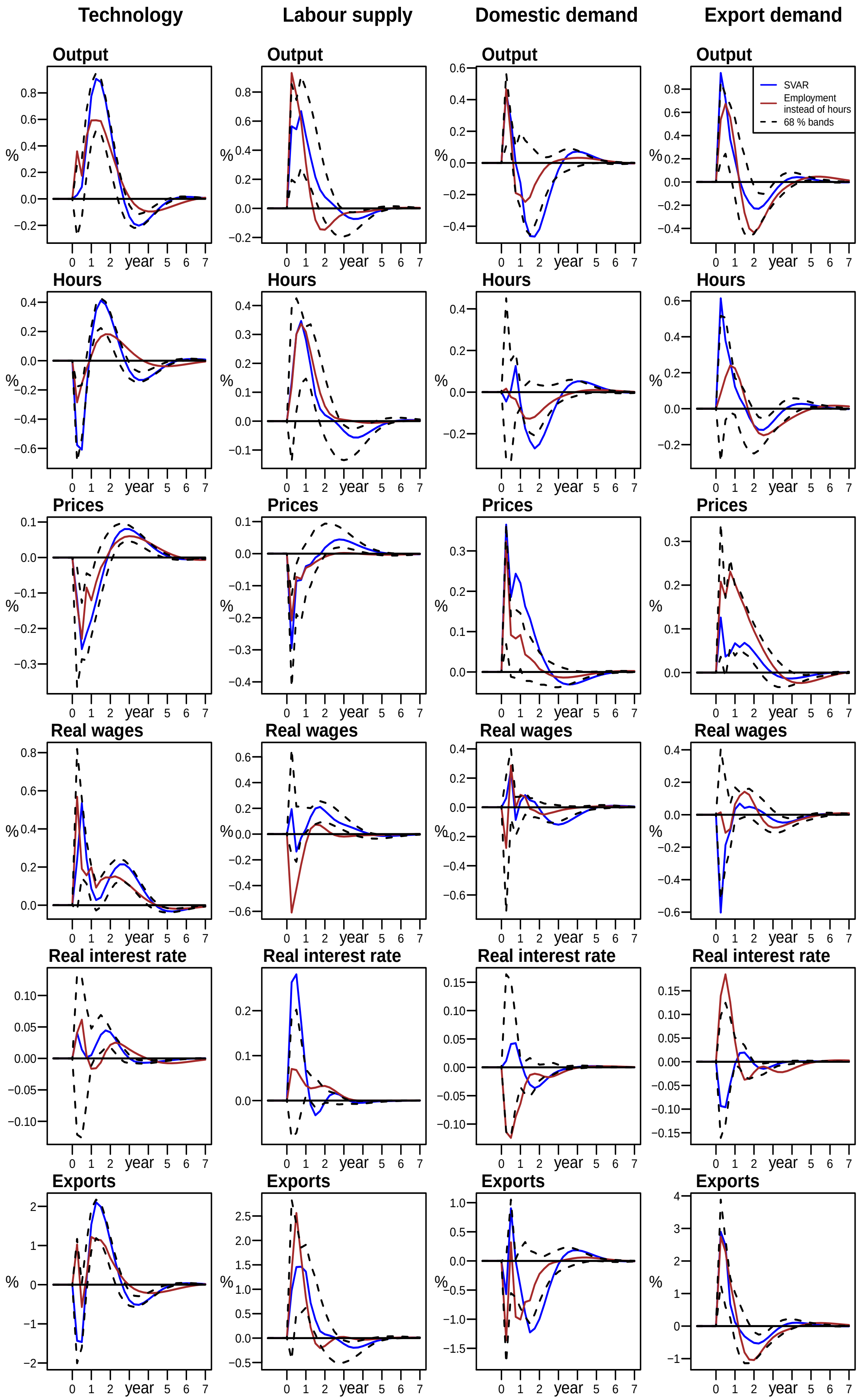

Figure 17: Robust check 12, employment instead of hours.

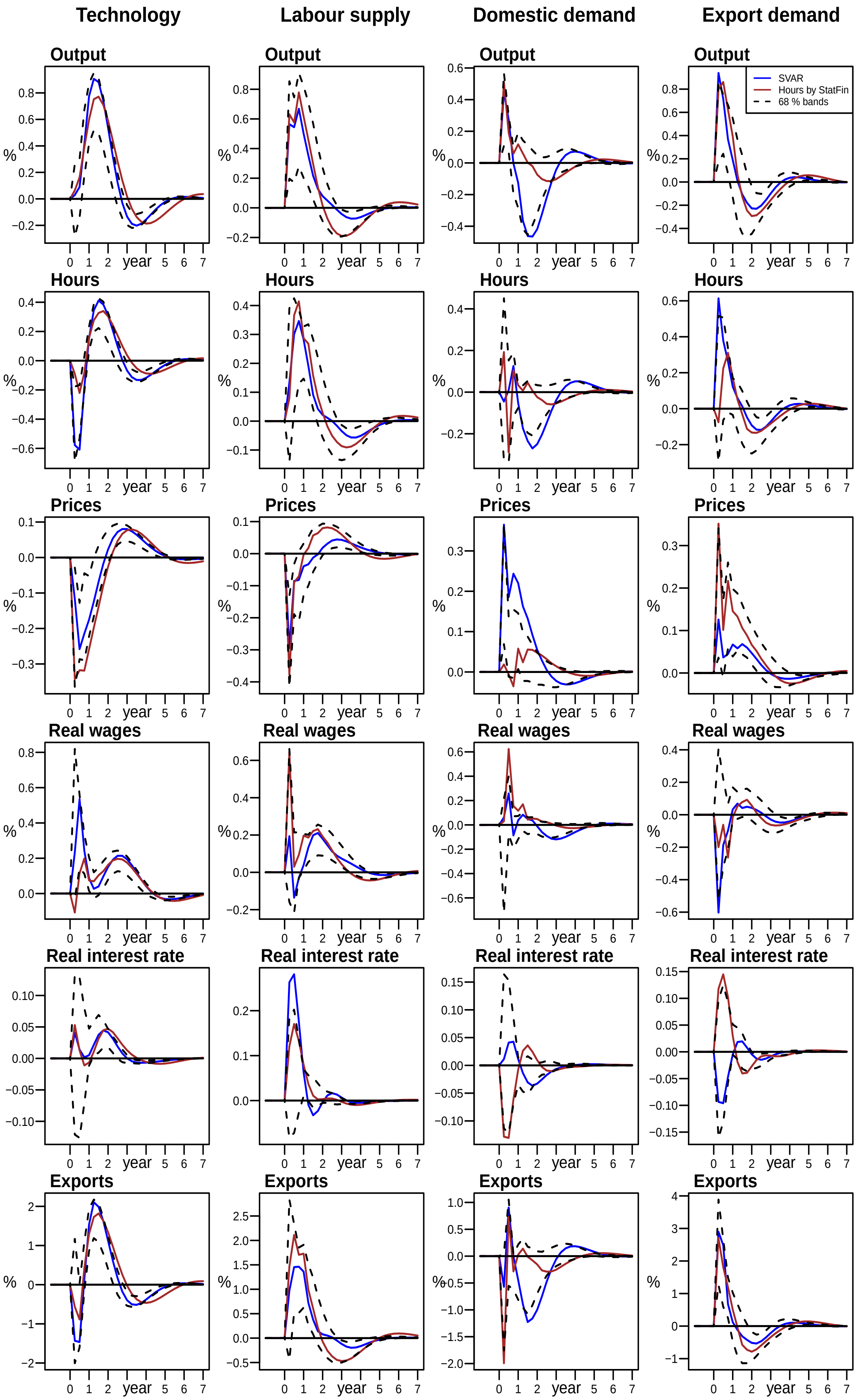

Figure 18: Robust check 13, hours seasonally adjusted by Statistics Finland.

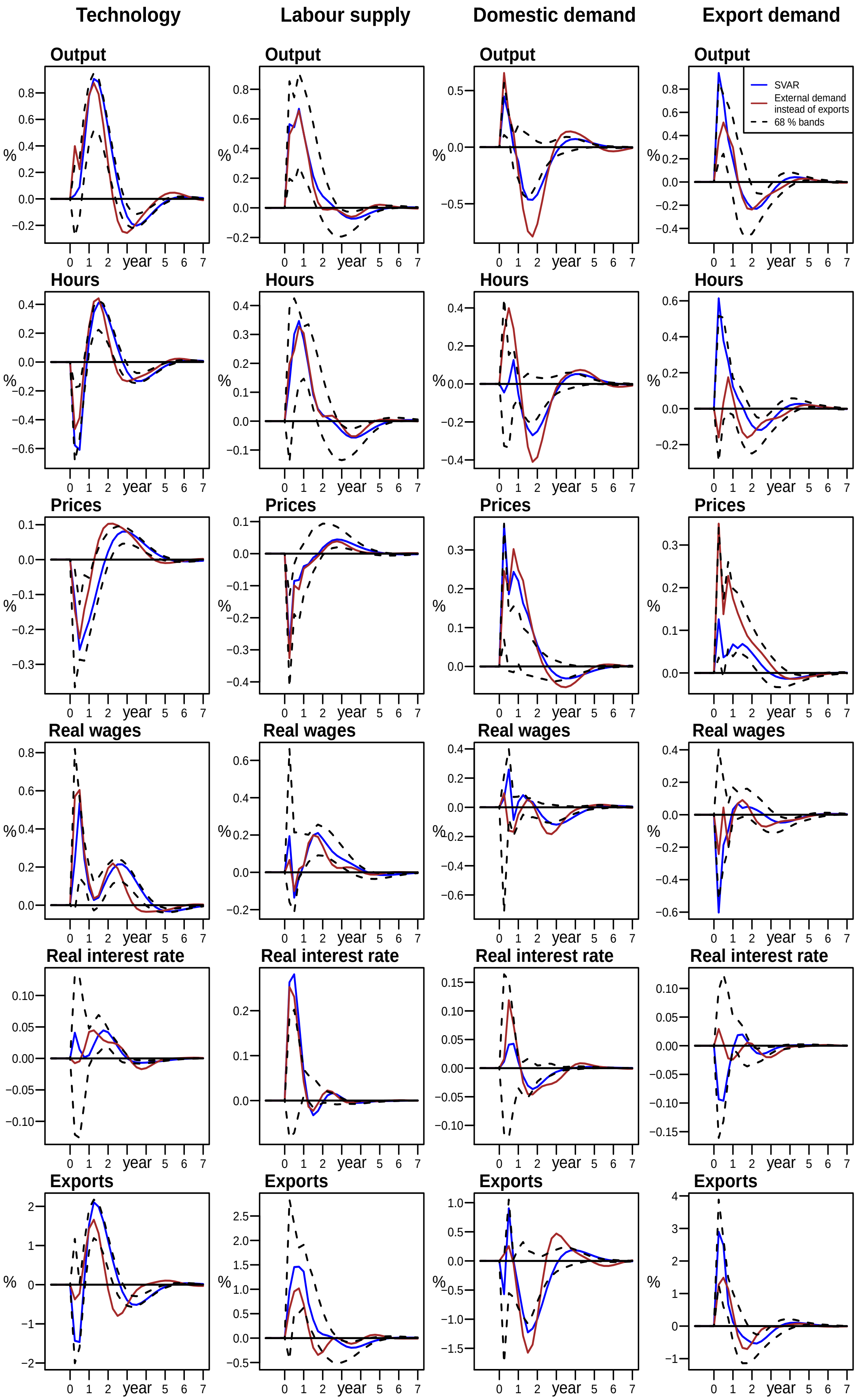

Figure 19: Robust check 14, external demand instead of exports.

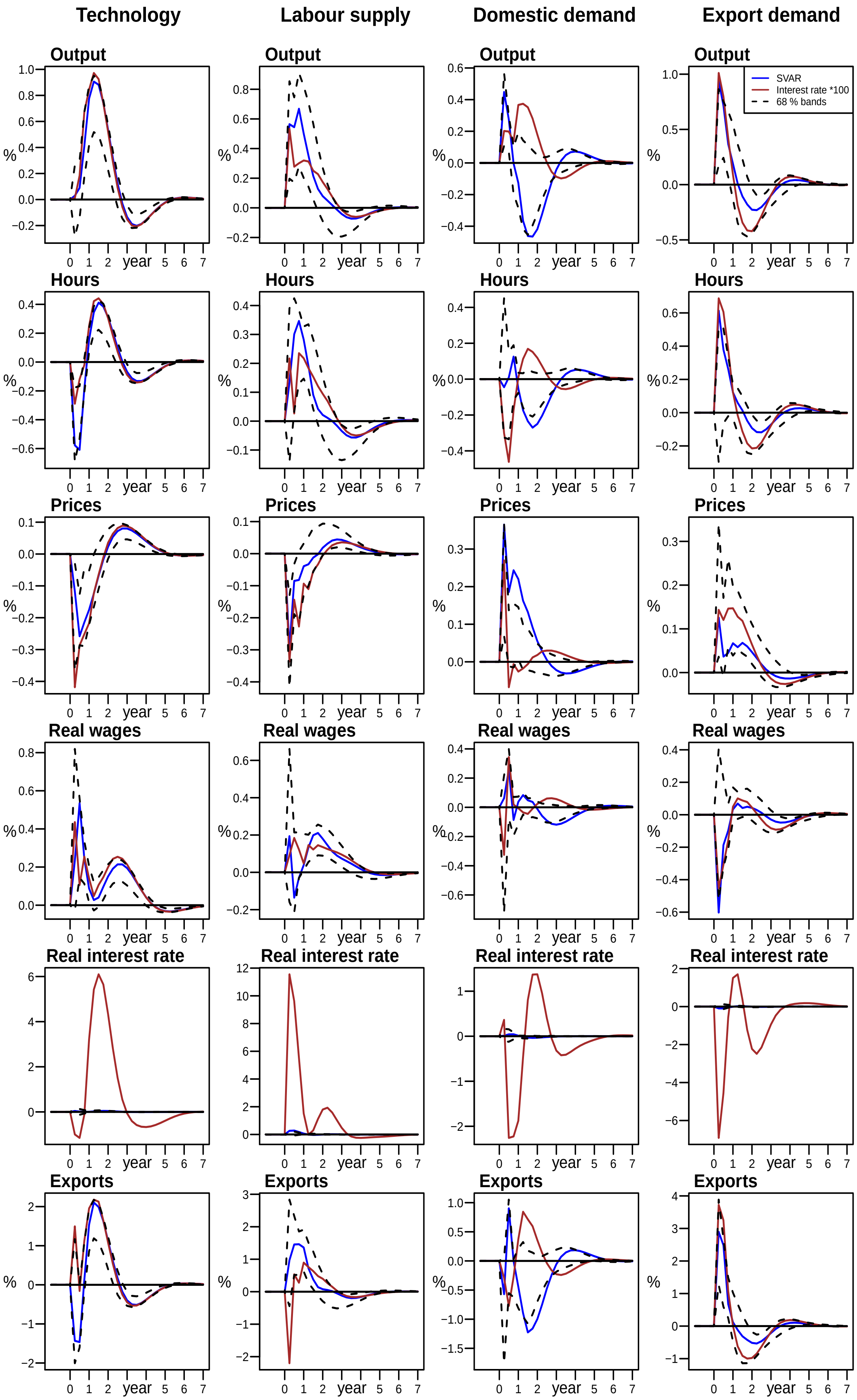

Figure 20: Robust check 15, real interest rate multiplied by 100.

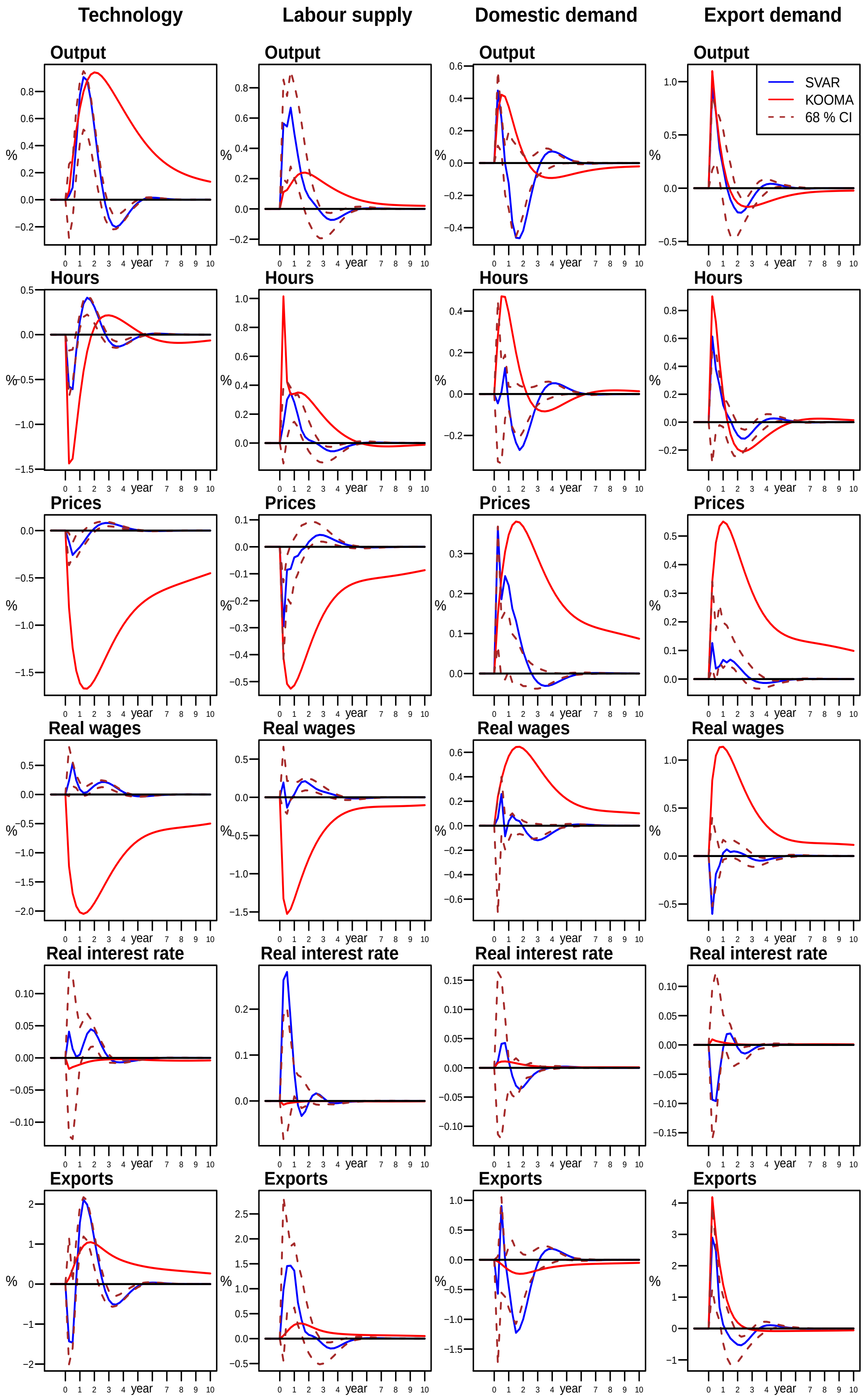

Figure 21: Comparison of impulse responses from KOOMA and SVAR models, with 68 % credible intervals.

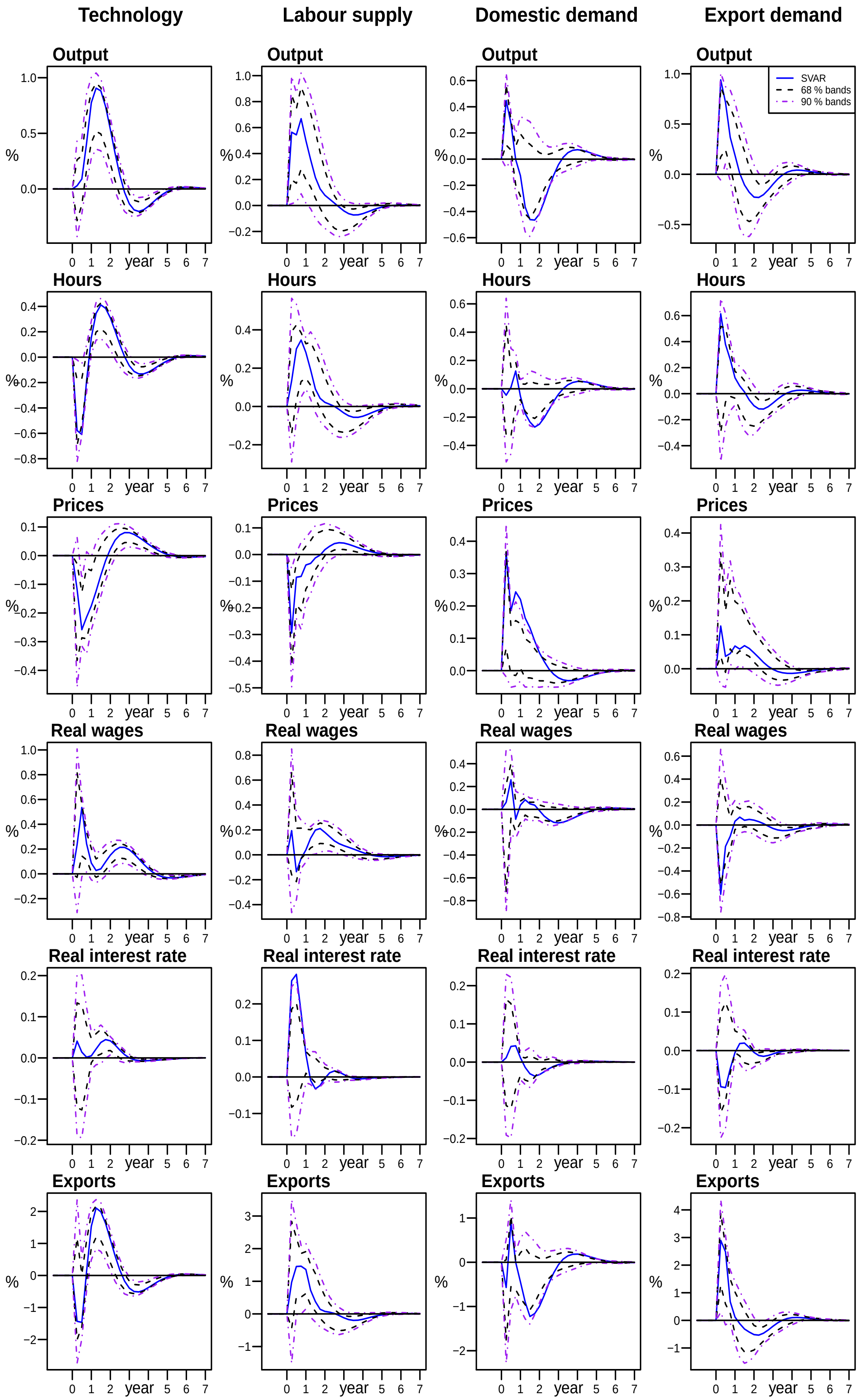

Figure 22: SVAR impulse responses with both 68 % and 90 % credible intervals.

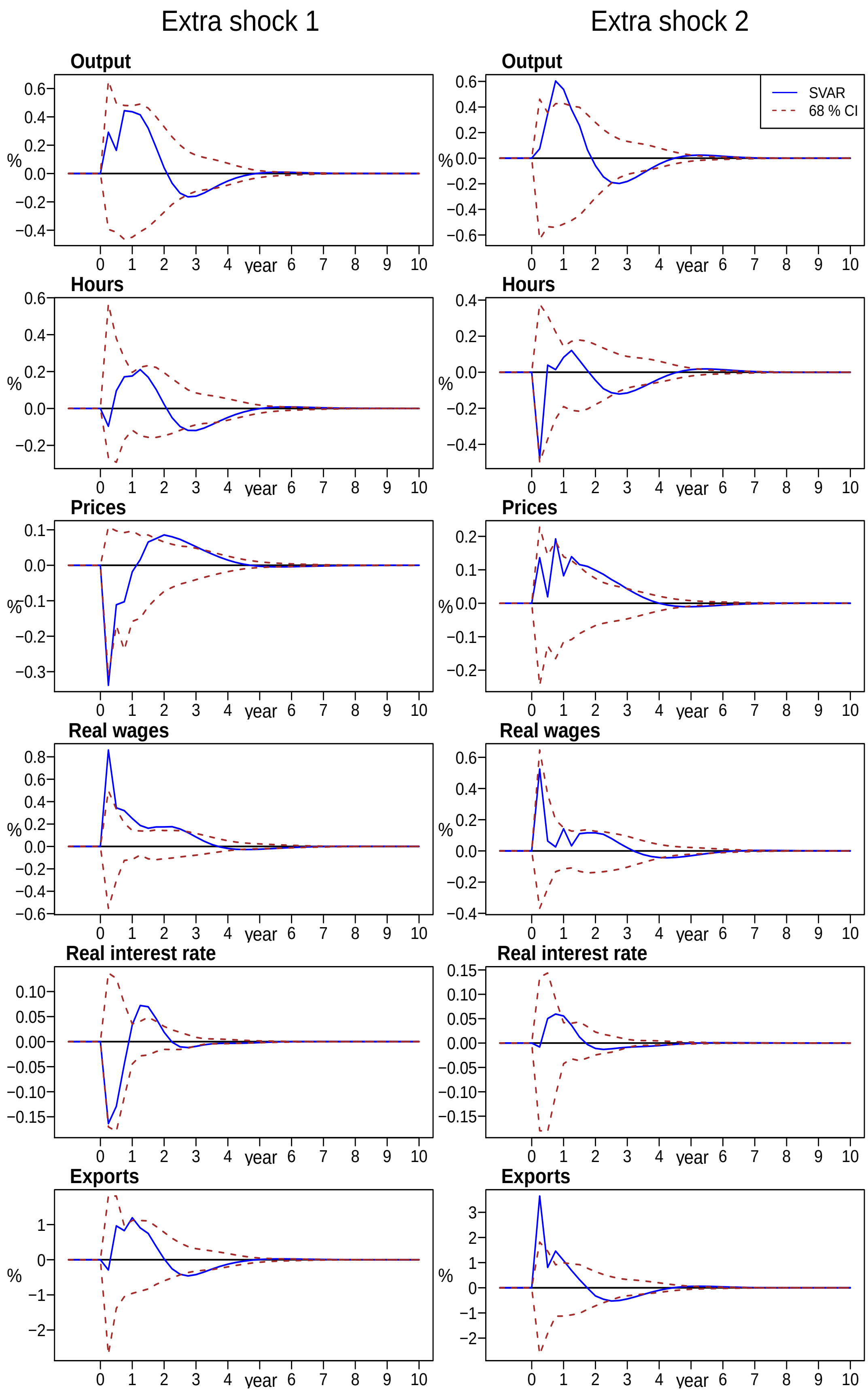

Figure 23: Impulse responses from two unidentified orthogonal shocks in SVAR model. These shocks are not addressed with economic interpretation, as they are linear combinations of the rest of the potential structural shocks influencing the variables.

Author:
Oliver Snellman

MINISTRY OF FINANCE
Snellmaninkatu 1 A
PO BOX 28, 00023 Government
Tel. +358 295 160 01
Fax 09 160 33123
www.financeministry.fi

ISSN 1797-9714 (pdf)
ISBN 978-952-367-044-0 (pdf)

October 2019